NC mosaic protein Blood transmiscibl HCV-S1 full-length

```
HCV capsid peptide
HCV capsid peptide
HCV capsid peptide
Non-A, non-B hepat
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Human hepatitis C
Human hepatitis C
HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                           Prototype peptide, HCV core protein p PT-NANB viral stru Hepatitis C virus Antigen pHCalol. Antigen pHCalol. HCV fragment 1 / I
                                                    Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Non-A, Non-B Hepat
Non-A, Non-B Hepat
Hepatitis C virus
CN14 fragment of H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NC mosaic protein amino acid fragment A.
                                                    AAR84559
AAR74562
AAR74562
AAR24946
AAR22510
AAR22256
AAW66487
AAR30689
                                                                                                                                                                                                                                                                              AAR29534
AAR29535
AAR29536
AAR29537
                                                                                                                                                                                                                                                                                                                                                                 AAR20770
AAR69545
AAY15425
AAW01865
AAR12597
AAR29160
AAR29161
                                                                                                                                                                                                                                                                                                                                       AAR29542
AAR25121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR41422
AAR41423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY06673 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17-JUN-1999 (first entry)
                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04-MAR-1999
AAY06673;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY0667
Search time 38.5455 Seconds (without alignments)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                 | SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseqp-embl/AA1982.DAT:*
| SIDSI/gcgdata/geneseqp-embl/AA1982.DAT:*
| SIDSI/gcgdata/geneseqp-embl/AA1982.DAT:*
| SIDSI/gcgdata/geneseqp-embl/AA1993.DAT:*
                                                                                                                                                                                                                                                                               1107863
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                           al number of hits satisfying chosen parameters:
                                                                                                                                                               151
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                  1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUMMARIES
                                                                                           August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                        A_Geneseq_19Jun03
                                                                                                                                                                                                                                                                                                          length: 0
length: 2000000000
                                                                                                                                                  US-09-491-146A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
Match Length
```

sed

Minimum DB Maximum DB

Database

Searched:

Perfect score: Sequence: Scoring table:

protein

ĕ

uo O

Peptide VIIIE base Anti-HCV antibody

fragment 1 / fragment 2 /

New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual

(USSH ) US DEPT HEALTH & HUMAN SERVICES

Khudyakov YE;

Fields HA,

WPI; 1999-204671/17.

HCV type 3 capsid Hepatitis C virus Encoded by Hepatit Encoded by Hepatit

AAB71258 AAR92968 AAR92971 AAR96547 AAR92970 AAR94473

20 23 23 11 11 14 14

28 189 191 191 191 319 191 470

100.0 100.0 100.0 100.0 100.0 99.3 97.4

151 151 151 151 151 151 151 147

mosaic protein

AAY06673

ΩI

DB

Score

Result Š.

Description

97US-0921887

25-AUG-1997;

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also proteided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against mosaic gene and protein is also useful for vaccination against suffection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, specially mosaic protein improve the sensitivity, specially mosaic protein improve the sensitivity, sequences AXY06673-683 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This invention describes a novel attenuated flavivirus live vaccine comprising a flavivirus mutant that has a deletion of at least 4 consecutive amino acids from the capsid protein, provided that the C-terminal hydrophobic region is not affected by the deletion. The vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Capsid protein; attenuated vaccine; virucide; antinflammatory; hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Score 151; DB 20;
100.0%; Pred. No. 1.5e-14;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNTIRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV type 3 capsid protein fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB71258 standard; protein; 189 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 2; 30pp; German.
                                  Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11-FEB-2002; 2002WO-AT00046.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-FEB-2001; 2001AT-0000272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-NOV-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match est Local Similarity luv...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              from the capsid protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Heinz FX, Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (HEIN/) HEINZ F X.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200266621-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29-AUG-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB71258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MAND/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB71258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     \overset{\times}{\times} \overset{\times}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

```
ö
wide range of flavivirus diseases, e.g. yellow fever, Japanese encephalitis, dengue, classical swine fever, bovine viral diarrhoea and hepatitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does not effect immune responses to important proteins; and (iii) can not generate a non-natural virus by recombination. The mutant viruses eliminate the need to produce large amounts of infectious/virulent viruses, and can be produced with less expense. The protective response to flavivirus lasts significantly longer than that to killed vaccines. This sequence represents a fragment of the capsid protein from Hepatitis c virus (HCV) type I described in the disclosure of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0%; Score 151; DB 17;
100.0%; Pred. No. 1.1e-13;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                    ; Score 151; DB 23;
; Pred. No. 1.1e-13;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus isolate HK10 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 PKPQKKTKKNTIRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEPT HEALTH & HUMAN SERVICES SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92968 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 4; Page 207; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                    100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95WO-US10398.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                              189 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAT16642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22-FEB-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SD (HSSD)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis.
                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR92968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR92968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
```

à g 

```
/note= "amino acid in this position is designated X in the specification, but codon usage shows that the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus; subtype; polymerase chain reaction; amplification;
PCR; primer; probe; antibody; infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                              DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR92936-R92987 are HCV core proteins derived from 52 different HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 151; DB 17;
Pred. No. 1.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 144..149 //label- Val, Ala, Glu, Gl;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'label- Met, Thr, Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US DEPT HEALTH & HUMAN SERVICES. US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /label- Val, Ala, Asp, Misc-difference 161
                                                                                                                                                                                                                                                                                                                                                                                                Claim 4; Page 209-210; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR96547 standard; peptide; 319 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label Met, Leu
                                                                                                                                                                                                                                                            Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 100.0%;
Best Local Similarity 100.0%;
Matches 28; Conservative 0;
                                                                                                                                     95WO-US10398
                                                                                                                                                                       94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /label- Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 156
                                                                                                                                                                                                                                                          Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                             WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 144
                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                               N-PSDB; AAT16645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
                                                                                                                                                                      15-AUG-1994;
                                                                  WO9605315-A2
                                                                                                                                     15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-MAR-1997
                                                                                                   22-FEB-1996
hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR96547;
                                                                                                                                                                                                        HSSU)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR96547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                             envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphcoytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Score 151; DB 17;
100.0%; Pred. No. 1.1e-13;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus isolate DK12 core protein.
                                                                                                                                                                                                                                          Hepatitis C virus isolate S52 core protein.
                                  PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (USSH ) US DEPT HEALTH & HUMAN SERVICES.
(USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR92971 standard; Protein; 191 AA.
                                                                                                                                     AAR92969 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 4; Page 208; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bukh J, Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                    95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94US-0290665
                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191 AA;
                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                    15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                        02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                  22-FEB-1996.
                                                                                                                                                                                                                                                                                               hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR92971;
                                                                                                                                                                       AAR92969;
                                                                                                                                                                                                                                                                           HCV; E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
                                                                                                   RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR92971
ID AAR
```

x cg

à qq XX DX DX XX

```
HCV; E1; envelope 1; core protein; HCV genctyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the produ. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   / Match 99.3%; Score 150; DB 17; Length 191; Local Similarity 96.4%; Pred. No. 1.6e-13; les 27; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                          Hepatitis C virus isolate S2 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Encoded by Hepatitis C virus clone JK3-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (USSH ) US DEPT HEALTH & HUMAN SERVICES (USSH ) US SEC DEPT HEALTH.
                     AAR92970 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR34473 standard; Protein; 470 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 4; Page 209; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                     95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                        94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1996-139709/14.
                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT16644
                                                                                                                                                                                                                                                                                     WO9605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                15-AUG-1995;
                                                                                                    02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-JUL-1993
                                                                                                                                                                                                                                                                                                                           22-FEB-1996
                                                                                                                                                                                                       hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                            AAR92970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR34473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
    AAR92970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR34473
                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The sequences AAR96526-R96578 represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes la-f, 2e-i, 2k, 2l, 3g, 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5' untranslated region (UR), the Core/El, NS4 or NS5B regions of the genome. This sequence represents amino acids 1-317 from the HCV types 7c and 8a isolated region (Cameroon and Vietnam, because of their Penelux countries, France, Cameroon and Vietnam, because of their abertant reactivities. The RNA was extracted, CDNA synthesised and PCR amplified, cloned and genotyped. The 5'UR, Core/El and NS5B regions were sequenced either directly or partially and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences can also be used to spentate the peptides AAR96424 R95524. The sequences can also be used to synthesise probes and primers for the cettion of HCV in a sample. The polypeptides can be used to detect anti-HCV antibodies, for HCV typing or to prevent HCV infections.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                      /note= "amino acid in this position is designated X in the specification, but codon usage shows that the only possible amino acid at this pos. is Gly"
only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%; Score 151; DB 17;
100.0%; Pred. No. 1.9e-13;
tive 0; Mismatches 0;
                                                                             'label≖ Val, Ala, Glu, Gly
                                                                                                                                                                                                                      label- Cys, Arg, Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                     label- Ser, Arg
                                                                                                                                                                                                                                                                                                                                           Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 25; Fig 3; 150pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      to prevent and treat infection
                                                                                                                                                                                                                                                            'label- Phe,
                                                                                                                                                                                                                                                                                                                                         'label- Met,
                                                                                                                                                                                                                                                                                                                                                                                   /label- Asn,
                                                                                                                                                                                                                                                                                                    'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95EP-0870076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95WO-EP04155.
                                                                                                               /label- Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Maertens G, Stuyver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1996-251460/25.
N-PSDB; AAT27958.
                  Misc-difference 167
                                                                                                                                                                                                   Misc-difference 172
                                                                                                                                                                                                                                         Misc-difference 174
                                                        Misc-difference 167
                                                                                                                                                                                                                                                                                                                    Misc-difference 232
                                                                                                                                                                                                                                                                                                                                                              Misc-difference 233
                                                                                                Misc-difference 17
                                                                                                                                                                                                                                                                               Misc-difference 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 319 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-JUN-1995;
21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                       WO9613590-A2
```

09-MAY-1996.

ö

Gaps

ö

JP05068562-A

23-MAR-1993

RESULT

ò a

38888888888

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-C. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                          Gaps
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 470;
 Length 470;
                                                                                                                                                                                                                                         HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA and cDNA of hepatitis C virus – useful as probes for diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97.4%; Score 147; DB 14;
llarity 96.4%; Pred. No. 1.1e-12;
Conservative 1; Mismatches 0;
Score 147; DB 14;
Pred. No. 1.1e-12;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Encoded by Hepatitis C virus clone JK3-C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                 28
                                                              PKPQRQTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NC mosaic protein amino acid fragment C.
                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 4; Page 30-32; 44pp; Japanese.
                                                                                                                                          Ź
                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ź
                                                                                                                                         AAR34475 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                         (SANW ) SANWA KAGAKU KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY06675 standard; Protein; 28
 Query Match 97.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                         91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1993-130638/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    See AAQ40425-Q40439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
nes 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAQ40433
                                                                                                                                                                                                                                                                                                       JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                         30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                 30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-JUN-1999
                                                                                                                                                                                         30-JUL-1993
                                                                                                                                                                                                                                                                                                                                23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY06675;
                                                                                                                                                                  AAR34475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY06675
                                                                                                                             AAR34475
                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                    à
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-B. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-Q40439.
                                                                                                                                                                                        cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-A. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                            Length 470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV; non-A, non-B hepatitis virus; NANBHV; liver disease;
                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV infection
                                                                                                                             DNA and cDNA of hepatitis C virus - useful as probes for
                                                                                                                                                                                                                                                                                            Score 147; DB 14;
Pred. No. 1.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polymerase chain reaction; diagnostic method
                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                           Encoded by Hepatitis C virus clone JK3-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 4; Page 28-30; 44pp; Japanese.
                                                                                                                                                               Claim 4; Page 26-28; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                AAR34474 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SANW ) SANWA KAGAKU KENKYUSHO
                                                             (SANW ) SANWA KAGAKU KENKYUSHO
                                                                                                                                                                                                                                                                                           Match 97.4%;
Local Similarity 96.4%;
les 27; Conservative 1
               91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91JP-0153736
                                       91JP-0153736
                                                                                                                                           diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1993-130638/16.
                                                                                        WPI; 1993-130638/16.
                                                                                                                                                                                                                                             See AAQ40425-Q40439
                                                                                                                                                                                                                                                                     470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAQ40432
                                                                                                       N-PSDB; AAQ40431
               30-MAY-1991;
                                        30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JP05068562-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR34474;
                                                                                                                                                                                                                                                                      Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                         AAR3447
                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                      ž
                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                               , xqxqxqxqqxqqxqx0000xx
0xqxqxqxqqxqqxqx00000x0
     ò
```

```
_{\rm G1y}
                                                                                                                         Len
                                                               /label= Ser, Arg,
                                                                                                                       label- Gln, Pro,
                                    label= Gly, Asp
                                                                                                                                                  label- Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gln
/label- Phe,
                                                                                                                                                                                                         abel- Asn,
                                                                                                                                                                                                                                                                 label- Arg,
                                                                                                                                                                                                                                                                                             abel- Leu,
                                                                                                                                                                                                                                                                                                                                                  abel= Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    abel- Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label= Glu,
                                                                                           'label Thr,
                                                                                                                                                                             abel- Gly,
                                                                                                                                                                                                                                     abel- Phe,
                                                                                                                                                                                                                                                                                                                                                                              label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                            label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label= Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      'label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'label= Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92JP-0051885
                                                                                                                                                                                                                                                                                                                                   Misc-difference 1036
                                                                                                                                                                                                                                                                                                                                                              Misc-difference 1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 2462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 2989
                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 1895
                                                                            Misc-difference 400
                                                                                                       Misc-difference 405
                                                                                                                                                              Misc-difference 418
                                                                                                                                                                                                                                                                             Misc-difference 759
                                                                                                                                                                                        Misc-difference 430
                                                  Misc-difference 398
                                                                                                                                  Misc-difference 41
                                                                                                                                                                                                                                                 Misc-difference 47
                                                                                                                                                                                                                    Misc-difference 43
                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                    Misc-difference
                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JP06105690-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-APR-1994
        ö
                                                                                                                                                                                                                                                                                                The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from notleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences ANYOGGN3-683 represent amino acid sequence of each monomer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NS5 region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                               New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96.7%; Score 146; DB 20;
96.4%; Pred. No. 7.9e-14;
Live 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                             (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      comprising the NC mosaic protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /label- Asp, Asn
Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label - His, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'label- Cys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /label- Leu, Phe
                                                                                                                                                                                                                                                                       Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Non-A, non-B hepatitis virus.
                                                                        98WO-US17385.
                                                                                                    97us-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JAN-1995 (first entry)
                                                                                                                                                        Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 96.4 ies 27; Conservative
                                                                                                                                                                                     WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28 AA;
                 WO9910506-A1
                                                                                                    25-AUG-1997;
                                                                        21-AUG-1998;
                                            04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR53417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
```

85666666666668

ò g Gaps

ö

Indels

Length 3010;

23;

```
N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3' untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a third polynucleotide region encoding a reporter molecule, flanked by first and second polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV-SI full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present sequence is that of an antigen used in a novel assay for determining anti-HCV antibody titre. A sample is diluted to an absorbance of more than 2.0 and the peptide antigen is added. The method is useful for monitoring the effect of interferon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Determination of the antibody titre against hepatitis C virus antigen - using dilute solution and HCV antigenic peptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; non-A, non-B hepatitis virus; antigen; immunoassay;
interferon treatment; monitoring; antibody titre; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 145; DB 16;
Pred. No. 1.6e-13;
                                                                                                                                                                                                                                                                                                                 Score 146; DB 2
Pred. No. 1e-11;
                                                                                                                                                                                                                                                                                                                                                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus core antigenic peptide.
                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 PKPORKTKRNTSRRPODVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     treatment on anti-HCV antibody titres.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 1; Page 10; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96.0%;
96.4%;
                                                                                                                                                                                                                                                                                                                 96.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94JP-0082160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94JP-0082160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR84559 standard; peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1995-386063/50.
                                                                                                                                                                                                                                                                          3010 AA;
                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP07260792-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JUN-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-OCT-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR84559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR84559
    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                          This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were Cl00 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on cDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The invention relates to nucleic acid construct which comprises an expression cassette including a first polynucleotide region including a 5' non-coding region (NCR) sequence of an RNA virus and at least an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                             Blood-transmissible non-A non-B hepatitis virus DNA - used for
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                         / Match 96.7%; Score 146; DB 15; Local Similarity 96.4%; Pred. No. 1e-11; nes 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (MOLE-) INST MOLECULAR & CELL BIOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                          Claim 1; Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAE20477 standard; Protein; 3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV-S1 full-length polyprotein.
                                                                                                                                                detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-JUL-2000; 2000US-220248P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ran YH, Lim SP, Lim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2002-280605/32.
                                                              WPI; 1994-163130/20.
                                                                                                                                                                                                                                                                                                                                                                                                                    3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EHRL/) EHRLICH G
                       (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAD33038
                                                                                    N-PSDB; AAQ63499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200208447-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31-JAN-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAE20477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
```

Matches

RESULT 13 AAE2047

g ò

ö

Gaps

ö

Length 40; Indels

RESULT 15

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR74561-63 and AAR74566 are respectively, 24mer, 40mer, 50mer and 30mer antigenic peptides derived from the hepatitis C virus clone S29 genes III and IV. They are useful in the detection of and diagnosis of Hepatitis C virus infection. The peptides are used in a detection method which has a higher sensitivity than conventional detection agents.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                    Agent for detection of hepatitis C - also hepatitis C virus (HCV) core protein antigen peptide(s), useful for immuno:detection of HCV \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96.0%; Score 145; DB 16; Length 40; 96.4%; Pred_.No. 1.6e-13; Live 0; Mismatches 1; Indels
                                                                                                                               Hepatitis C virus; detection; immunodetection; sensitive.
                                                                                                  Hepatitis C virus core protein antigen peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNTIRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 31
               AAR74562 standard; peptide; 40 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 1; Page 7; 10pp; Japanese.
                                                                                                                                                                                                                                                    93JP-0273938.
                                                                                                                                                                                                                                                                                 93JP-027393B
                                                                   04-JAN-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                            (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
                                                                                                                                                                                                                                                                                                                                                       WPI; 1995-175362/23.
                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40 AA;
                                                                                                                                                                                         JP07097398-A.
                                                                                                                                                                                                                                                    28-SEP-1993;
                                                                                                                                                                                                                                                                               28-SEP-1993;
                                                                                                                                                                                                                      11-APR-1995.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                          AAR74562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
AAR74562
XX
AC AAR7
XX
DT 04-J
XX
XX
XX
XX
KW HEP8
XX
XX
PD HEP8
XX
PN JP0;
XX
PD 11-1
                                                                                                                                                                                                                                                                                                            ò
```

search completed: August 7, 2003, 11:14:05
Job time : 39.6364 secs

돈

```
Genome polyprotein - hepatitis C virus (genotype 3, NI) (fragment)
N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 3, NI
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accession: $41358
R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: $41341
A.Accession: $41358
A.Molecule type: genomic RNA
A.Residues: 1-114 < VANA
A.Residues: 1-114 < VANA
A.Residues: 1-114 < VANA
A.Residues: Caprus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
F.1-114/Product: core protein #status predicted < CMAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  lyprotein
lyprotein
lyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      yprotein
yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           rprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                 August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       зепоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                   283308
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                                                                                                                   .al number of hits satisfying chosen parameters:
                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                        283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S41357
S41378
S41371
S41371
S41370
S41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S41358
S41360
PC2061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S41288
S41361
S19875
S41356
PC2060
                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S41359
                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                         US-09-491-146A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123
782
108
411
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Match Length
                                                                                                                                                                                                                                                                                                                                                                                             PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                      pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0
97.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0
                                                                                                                                                    Perfect score:
Sequence:
                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                                                                                         searched:
                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ဍ
```

Gaps

ö

Indels

ö

0; Mismatches

28

Length 114;

DB 2;

Score 151;

100.0%;

genome polyprotein genome polyprotein genome polyprotein hepa polyprotein hepa polyprotein hepa genome polyprotein	NTS of y	No. 2.4e-14;
S21471 S12707 PC1284 J01925 J10712 S18031 S18031 S18032 PN0677 JQ0883 JQ0883 JQ0883 S18032 S18032 S18032 S18032 S18032 S18032 S18033 S18032 S18033 S18033	ALIGNMENTS  tis C virus (genoty s ce_revision 26-Jul- G.E.M.; Brouwer, J. Library, January 19 hepatitis C virus g hepatitis C virus g A A A A A A A A A A A A A A A A A A A	Pred.
	us nce_revi G.E.M.; Library hepatit hepatit hotype 3, otype 3, core F ein *sta	; 6
. 3469 520 520 520 523 782 782 782 782 782 3010 3010 3010	rotein - hepatitis C virus core protein epatitis C virus enotype 3, N2 enotype 3, N2 extisson S4139 L.J.; Kleter, G.E.M.; Brou L.J.; Kleter, G.E.M.; Brou it he EMBL Data Library, Jan n: Analysis of hepatitis C number: S41341 sype: genomic RNA 1-114 <-VAN> 1-114 <-VAN> 1-114 <-VAN> 341359 ype: genomic RNA 1-114 <-VAN> 1-114 <-VAN 1	100.08;
00000000000000000000000000000000000000	String String #3, Ce	Similarity
	polyprotein ins: core press: hepatit. ity: genotype is penotype is penotype is penotype is penotype is penotype is penotype is penotype; penotype is penotype; penotype is penotype; penot	cal Sın
@@@@@@@@@@############################	RESULT 1 S41359 genome polyprotein - hepatitis C N:Contains: core protein C; Species: hepatitis C virus A; Variety: genotype 3, N2 C; Date: 19-May-1994 #sequence_rev C; Accession: S41359 R; Van Doorn, L.J.; Kleter, G.E.M. submitted to the EMBL Data Librar A; Description: Analysis of hepati A; Reference number: S41341 A; Accession: S41359 A; CSS-references: EMBL:229462; A; Experimental source: genotype C; Superfamily: hepatitis C virus C; Keywords: capsid protein; core F;1-114/Product: core protein #st	Best Local

us-09-491-146a-23.rpr

~

ö

Gaps

ö

Indels

Length 492;

```
R; Seelig, R.
submitted to the EMBL Data Library, December 1993
A; Reference number: $41288
A; Accession: $41288
A; Accession: $41288
A; Molecule type: genomic RNA
A; Rossidues: 1-492 < SEED.
A; Rossi-references: EMBL:X76918
A; Cossi-references: EMBL:X76918
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; envelope protein; nonstructural protein;
E; 1-191/Product: core protein #status predicted <COR>
E; 372-472/Product: envelope protein #status predicted <NSI>
E; 373-492/Product: NSI protein (fragment) #status predicted <NSI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A:Variety: isolate JK3
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C;Accession: S19875
C;Accession: S19875
S;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
S;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
S;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
S;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
S;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
S;Hoscription: Squence analysis of putative structural regions of Hepatitis C Viru.
A;Residues: Squence: Square Saneko, S;Residues: 1-782 **GHONA**
A;Residues: 1-782 **GHONA**
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 prote
C;Species: hepatitis C virus
.Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
.Accession: S41288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C. Accession: S41361
R. van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A. Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A. Beference number: S41341
A. Accession: S41361
A. A. Molecule type: genomic RNA
A. Residues: 1-123 cVAN>
A. Cross-references: EMBL: Z29464; NID: 9443890; PIDN: CAA82602.1; PID: 9443891
A. Experimental source: genomype 3, NA
A. Cross-references: EMBL: Z29464; NID: 9497protein
C. Superfamily: hepatitis C virus genome polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
C. Keywords: capsid protein; core protein; polyprotein
F:1-123/Product: core protein #status predicted cMAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 3, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
100.0%; Score 151; DB 2;
Best Local Similarity 100.0%; Pred. No. 9.9e-14;
Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97.4%; Score 147; DB 2; 92.9%; Pred. No. 9.5e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Variety: genotype 3, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N:Contains: core protein
C:Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
nes 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       olecule type: mRNA
h, wesdques: 1-411 <LLID>
A; Cross-references: CB:L12355; NID:g410169; PIDN:AAA20155.1; PID:g410170
C; Superfamily: hepatitis c virus genome polyprotein
C; Superfamily: hepatitis c virus genome polyprotein
C; Superfamily: hepatitis c virus genome polyprotein
C; Superfamily: hepatitis c virus genome polyprotein; glycoprotein; nonstructural protein;
F; 192-383/Product: envelope protein El *status predicted <SPE>
F; 384-411/Product: nonstructural protein El*(NSI) *status predicted <NPE>
F; 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Propose polyprotein N2 - hepatitis C virus
N; Contains: envelope protein E1; nonstructural protein E2/NS1
C; Species: hepatitis C virus
C; Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
C; Accession: PC2061
R; Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A; Tille: Identification of the third major genotype of hepatitis C virus in France.
A; Reference number: PC2060; MUID:94197744; PMID:8147893
                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                    RESULT 3
S41360
genome polyprotein - hepatitis C virus (genotype 3, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Varitety: genotype 3, N3
C;Accession: S41360
C;Accession: S41360
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A,Cross-references: EMBL:229463; NID:9443889; PIDN:CAA82601.1; PID:9443889
A,Experimental source: genotype 3, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-124/Product: core protein #status predicted <MAT>
                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 124;
                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NSI protein
C;Species: hepatitis C virus
                   Pred. No. 2.4e-14;
Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Score 151; DB 2;
100.0%; Pred. No. 2.6e-14;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 8.3e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 151;
100.0%; Pred. No. 8
:ive 0; Mismatche
                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                Best Local Similarity 100.0%; F
Matches 28; Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 100.(
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Olecule type: genomic RNA .esidues: 1-124 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Accession: S41360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
S41288
                                                                                                                            ô
                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
```

ö

Gaps

.; 0

us-09-491-146a-23.rpr

à g

carr

```
A; Status: preliminary
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-3010 cTAN
A; Residues: 1-3010 cTAN
A; Residues: 1-3010 cTAN
A; Cross-references: GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g22161
A; Experimental source: HCV-JT
A; Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: APP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; F; 2-115/Product: capsid protein M #status predicted cAPP.
F; 116-191/Product: major envelope protein B #status predicted cARE>
F; 390-729/Product: major envelope protein NSI #status predicted cNSI>
F; 390-1006/Product: nonstructural protein NSI #status predicted cNSI>
F; 120-1317/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif B (F-loop)
F; 1312-1317/Region: nucleotide-binding motif B (F-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein S29 (core protein region) - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Species: S21336
R;Sato, A.
R;Sato, A.
R;Sato, A.
R;Sato, A.
R;Sato, A.
R;Setence number: S21336
A;Reference number: S21336
                                                                                 genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nons protein NS4a; nonstructural protein NS5scies: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                     R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Virus Res. 23, 39-53, 1992
Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese
A;Reference number: A45573; MUID:92295714; PMID:1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NID: 959492; PIDN: CAA46517.1; PID: 959493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F:2014-3010/Product: nonstructural protein NS5 #status predicted <NAS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment) N;Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match 96.7%; Score 146; DB 1; Local Similarity 96.4%; Pred. No. 3.1e-12; les 27; Conservative 0; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 145; DB 2;
Pred. No. 1.3e-13;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTKRNTIRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 96.0%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Cross-references: EMBL:X65548;
C;Superfamily: hepatitis C virus
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         procession: protein N1 - hepatitis C virus

Contains: envelope protein E1; nonstructural protein E2/NS1

Contains: envelope protein E1; nonstructural protein E2/NS1

Species: hepatitis C virus

Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000

C; Accession: PC2060

R; Li, J.S.; Vitvitiski, L.; Tong, S.P.; Trepo, C.

Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994

A:Title: Identification of the third major genotype of hepatitis C virus in France.

A; Reference number: PC2060

                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
F;1-191/Product: core protein #status predicted <MAT1>
F;192-383/Product: envelope protein 1 #status predicted <MAT2>
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
N;Contains: core protein
C:Species: hepatitis C virus
'ariety: genotype 2, N5
ate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIPA
                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   c, Accession: S41356
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by 3
A; Reference number: S41341
                                                                                                                                                                                                                                Score 147; DB 2; Length 782;
Pred. No. 5.8e-13;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 146; DB 2; Length 411;
Pred. No. 4.38-13;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A/Cross-references: EMBL:229459
A/Experimental source: genotype 2, N5
C/Superfamily: hepatitis C virus genome polyprotein
C/Keywords: capsid protein; core protein; polyprotein
F/1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 146; DB 2;
Pred. No. 1.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKPOROTKRNTIRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                         97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           / Match 96.7%;
Local Similarity 96.4%;
les 27; Conservative
                                                                                                                                                                                                                                                                                    Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S41356
A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
```

g

ð

ö

ö

QQ

```
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
C; Species: hepatitis C virus
A; Variety: genotype 2, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41353
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Accession: S41353
A; Molecule type: genomic RNA
A; Residues: 1-108 < VAN>
A; Residues: 1-108 < VAN>
A; Reperimental source: genotype 2, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted < MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
S41357
genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41357
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C, Accession: $41355
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: $41341
A; Accession: $41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA A; Reference number: $41341 A; Accession: $41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95.4%; Score 144; DB 2; Length 108; 96.4%; Pred. No. 2.2e-13; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95.4%; Score 144; DB 2; Length 108; 96.4%; Pred. No. 2.2e-13; ive 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Experimental source: genotype 2, N4
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKTKRNTIRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Cross-references: EMBL: 229460
A; Experimental source: genotype 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A, Molecule type: genomic RNA
A, Residues: 1-108 <VAN>
A, Cross-references: EMBL: 229458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 95.4
Best Local Similarity 96.4
tches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Molecule type: genomic RNA Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
ö
                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                  N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C; Accession: $41348
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A; Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                             - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                   Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 95.4%; Score 144; DB 2; L
Best Local Similarity 96.4%; Pred. No. 2.2e-13;
Matches 27; Conservative 0; Mismatches 1;
C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capsid protein; core protein; polyprotein F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A. Experimental source: GMBL:229451
A. Experimental source: genotype 1, N6
C. Superfamily: hepatitis C virus genome polyprotein
C. Keywords: cappid protein; core protein; polyprotein
F. 1-108/Product: core protein #status predicted < MAT>
                                                                                                   Score 144; DB 2;
Pred. No. 2.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                 7, 2003, 11:21:46
                                                                                             Query Match 95.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: August
Job time : 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41348
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28 US-09-491-146A-23 151

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

al number of hits satisfying chosen parameters:

127863 seqs, 47026705 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	ript	Q00269 h genome po	hepatitis			P27960 hepatitis c			h genome		h genome	c	h genome	h genome	brugia ma	P09958 homo sapien		080361 nicotiana t					P50894 fugu rubrip		mus m	015530 homo sapien	_	O55173 rattus norv	Q28193 bos taurus	P51816 homo sapien		$\sim$	Pl0637 mus musculu	P19332 rattus norv
SUMMARIES	OI .	POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	SYN_BRUMA	FURI_HUMAN	POLG_BVDVŅ	RK4_TOBAC	CAT3_MAIZE	FN3X_HUMAN	Y401_BUCAP	RL29_YEAST	RS7_FUGRU	RS7_HUMAN	FURI_MOUSE	PDPK_HUMAN	PDPK_MOUSE	PDPK_RAT	FURI_BOVIN	FMR2_HUMAN	YB1_XENLA	YAB1_YEAST	TAU_MOUSE	TAU_RAT
	DB	-	٦	-	٦	П	-	П	-	-	-	-	-	Н	-	٦	٦	~	7	٦		7			. ,					٦	7		7	7
	Length	3010	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3011	548	794	3988	282	496	309	219	28	194	194	793	556	559	559	797	1311	303	638	732	751
œ	Query Match	96.7	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4			90.1					32.5			31.1					30.5		。	。	٥.	30.1	ö	30.1
	Score	146	144	144	144	144	144	144	144	144	144	144	137	136	54	S	50.5	49	4	48.5	84	47	47	47	47					4	ŝ	45.5		•
	Result No.	п	7	m	4	S	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	52	56	27	28	58	30	31	32	33

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its mes by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).

EMBL; D11168; BAA01943.1; -.

InterPro; IPR001410; DEAD.

PIR; A45573; A45573. PDB; 1A10; 25-WAR-98 PDB; 1JYP; H-JAN-98. MEMOPS; S29.001; -. MEROPS; U39.001; -.

Q9z0e2 mus muscul	049937 spinacia ol	P37271 arabidopsis	P21525 drosophila	091713 xenopus lae	095405 homo sapien	P47390 mycoplasma	027036 methanobact	P08855 oryctolagus	P48155 manduca sex	P02362 xenopus lae	Q9czx5 mus musculu
CHRD_MOUSE	RK4_SPIOL	PSY_ARATH	FRA_DROME	CHRD_XENLA	MADI_HUMAN	Y144_MYCGE	VATA_METTH	ICAL_RABIT	RS7_MANSE	RS7_XENLA	PII1_MOUSE
-	н	٦	-	-	٦	-	٦	Н	ч	-	ч
948	293	422	595	941	1425	279	584	718	190	194	332
	_	80	æ.	æ.	æ	'n.	'n	٥.	٦.	۲.	۲.
30.1	29.8	59	53	53	29	29	29	5	53	5	53
	45 29.8				-						44 29

## ALIGNMENTS

g

```
CELLULAR ANINOPERIDAES.
CAPSID PROTEIN C POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART; SM0487; DEXDC: 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 146; DB 1; Length 3010;
Pred. No. 2.3e-13;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94A1C77435D642BB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
                                                             InterPro: IPR002531; HCV_NS1.
InterPro: IPR002518; HCV_NS2.
InterPro: IPR00146; HCV_NS3.
InterPro: IPR00146; HCV_NS4.
InterPro: IPR00146; HCV_NS4.
InterPro: IPR00140; HCV_NS4.
InterPro: IPR00156; HCV_NS3.
InterPro: IPR00106; HCV_RGRP.
InterPro: IPR001095; RNA_pol_DS_PS.
InterPro: IPR00166; HCV_NS3; I.
IPR ms. PF01066; HCV_NS3; I.
IPR ms. PF01006; HCV_NS4; I.
IPR ms. PF01006; HCV_NS5; I.
IPR ms. PF01006; HCV_NS5; I.
IPR ms. PF0100998; VITAL_RGRP; I.
IPR ms. PF000091; HCV_NS1; I.
IPR ms. PF000091; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       96.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        540
556
576
623
645
2041
2240
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                           InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
STATE FOR STATE OF ST
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-92230232; PubMed-1314459;
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
A Okamoto H., Fuxuai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fuxuai K., Okada S.I., Yamamoto S.;
Tanaka T., Fuxuai S., Tsuda F., Mishiro S.;
Tanaka T., Fuxuai S., Tsuda F., Mishiro S.;
Tanaka T., Fuxuai S., Tsuda S., Tsuda Batitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
Tenotypes.";
Tenot
POLG_HCVJ2 STANDARD; PRT; 513 AA.
PDT959;
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
17-OCT-2001 (Rel. 40, Last annotation update)
18-OCT-2001 (Rel. 40, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Þ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTEWTIAL).
MATRIX PROTEIN (POTEWTIAL).
MATRIX ENVELOPE PROTEIN E (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTEWTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . .) (POTENTIAL) . . . (POTENTIAL) . . . (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC, N-LINKED) (GLCNAC, N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV.capsid.
InterPro; IPR002521; HCV.capsid.
InterPro; IPR002521; HCV.capsid.
InterPro; IPR002531; HCV.car.
InterPro; IPR002531; HCV.cap.
Pfam; PF01543; HCV.capsid; 1.
Pfam; PF01542; HCV.capsid; 1.
Pfam; PF01542; HCV.cor; 1.
Pfam; PF01540; HCV.cor; 1.
Pfam; PF01560; HCV.ens; 1.
Probom; PB168662; HCV.ens; 1.
Polyprotein; Glycoprotein; Coat pri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10074; BAA00968.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
```

ö

ö

Gaps

```
NCBI_TaxID=31644;
                                                                                                                             POLG_HCVHK
Q01403;
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
                                                                                                           RESULT 4
POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                               ò
                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMIS custation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                    Gaps
                                                                                                                                   01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GENVelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein El (GP68) (GP69) (GP60) (NS1)] (Fragment).
                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                  Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                    1; Indels
943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
                  Score 144; DB 1;
Pred. No. 6.6e-14;
                                                                                                                   520 AA.
                                  0; Mismatches
                                                                                                                                                                                          Hepatitis C virus (isolate HCV-476) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BY SIMILARITY
                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                               Transmembrane; Nonstructural protein.
                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_core.
InterPro; IPR002531; HCV_senv.
Efam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D10688; BAA01530.1; -.
                 Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                   STANDARD;
513 AA;
                                                                                                                                                                                                                      NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1116
3384
3384
2209
3305
4418
4418
4418
                                                                                                                                                                                                             Hepacivirus
                                                                                                                  POLG_HCVH4
Q01404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
တ္တ
                                                                                                                  a
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR ANINOPEPIDAEL
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (CPOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. ...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GROBD (GP70) (NS1)] (Fragment)
Hepatitis C virus (1solate HCV-KF) (HCV).
Hepatitis C virus (1solate HCV-KF) (HCV).
                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Abe K., Inchauspe G., Fullsawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                             ö
                                                                                     Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C
520 520
520 AA; 56499 MW; AA135246CF20D525 CRC64;
                                                                                     Score 144; DB 1;
Pred. No. 6.7e-14;
                                                                                                                                                                                                                                                                                                                                                                                                520 AA
                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; JQ1925; JQ1925.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_core.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_env.
InterPro; IPR00251; HCV_env.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01560; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-93019030; PubMed-1383400;
                                                                            95.4%;
milarity 96.4%;
Conservative 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   115
191
383
369
209
233
234
418
                                                                               Query Match
Best Local Similarity
Matches 27; Conserv
```

us-09-491-146a-23.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ***MEDLINE-9220232; PubMed-1314459;
***A MEDLINE-9220232; PubMed-1314459;
***A Chance Ch. Kutai K., Okada S.I., Yamamoto K., Lizuka H.,
**A Tanaka T., Fukuda S., Tsuda F., Mishino S.;
**A Tanaka T., Fukuda S., Tsuda F., Mishino S.;
**A Tanaka T., Fukuda S., Tsuda F., Mishino S.;
**A Tanaka T., Fukuda S., Tsuda F., Mishino S.;
**A Tanaka T., Fukuda S., Tsuda F., Mishino S.;
**A Tanaka T., Fukuda S., Mishino S.;
**A Tanaka T., Fukuda F., Mishino S., Mishino S., Mishino T., Mi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
Protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                    Gaps
     N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein.
REMOVED FROM CAPSID PROTEIN C BY
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                          Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C
                                                                                                                            Score 144; DB 1; Length 52.
Pred. No. 6.7e-14;
424 N-LINKED (GLCNAC. . .) (PC
431 N-LINKED (GLCNAC. . .) (PC
449 N-LINKED (GLCNAC. . .) (PC
520 520 56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                           737 AA
                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                       EMBL, D10075; BAA00969.1; -
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
                                                                                                                                                                                  96.48;
                                                                                                                                                          95.48;
                                                                                                                            Query Match
Best Local Similarity 90.4.
Best Local Similarity 90.4.
                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
     424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVJ5
                             CARBOHYD
CARBOHYD
NON_TER
SEQUENCE
     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
POLG_HCVJ5
ID POLG_H
AC P27960
DT 01-AUG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
     당당당당
                                                                                                                                                                                                                                                                                                         g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Æ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL 3NA REPLICATION.
-!- SUBBNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A MUD GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-92230232; PubMed-1314459; Okamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Chanaka T., Fukuda F., Mishiro S., Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                Length 737;
                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                   81207 MW; 3AF699D82AD501B1 CRC64;
                                                                                                                                                                                                                                                             Score 144; DB 1;
Pred. No. 9.8e-14;
0; Mismatches 1;
                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                           (GLCNAC.
                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                   (GLCNAC.
                                                                                             (GLCNAC.
                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                     2
                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
                                                                N-LINKED
N-LINKED
N-LINKED
                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro: IPR002522; HCV_capsid.
Interpro: IPR002521; HCV_core.
Interpro: IPR002519; HCV_core.
Interpro: IPR002513; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                               95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                    737
737 AA;
                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genotypes.
                                                                                                                                                                                                     CARBOHYD
NON_TER
SEQUENCE
            CARBOHYD
                                                                                             CARBOHYD
                                                                                                         CARBOHYD
                                                                                                                       CARBOHYD
                                        CARBOHYD
                                                                  CARBOHYD
                                                                               CARBOHYD
                                                                                                                                    CARBOHYD
                                                                                                                                                   CARBOHYD
                                                                                                                                                                CARBOHYD
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                       RESULT
ద
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-1992 (Rel. 42, Last annotation update)
67-1992 (Rel. 42, Last annotation update)
67-1902 (Gontains: Capsid protein C (Core protein) (P22);
67-1903 (GP70) (NS1); Protein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
6EC 3.4.22.-); Proteases/helicase NS3 (P70) (Heperitiin)
6EC 3.4.21-98); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); Nonstructural protein NS5A (P670) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate BK) (HCV).
                                                                                                                CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                      protein.
REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 144; DB 1; Length 737; Pred. No. 9.8e-14;
                                                                      protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             737
81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                    N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
ProDom; PD18662; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat
Transmembrane; Nonstructural pr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 96.4%;
nes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVBK
                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                     INIT_MET
                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P26663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
POLG_HCVBK
                                                                                                                                     CHAIN
                                                                                                                                                                      CHAIN
                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                                                                                                                                                **RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

**MEDLINE-9827846; PubMed-9568891;

**A Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,

**Steinkuehler C., Tomeil.L., de Francesco R., Kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Tomplex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Tomplex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Tomplex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Tomplex of NS3 protease and NS4A peptide of BK strain hepatitis C

"TOMPORPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RAN REPLICATION.

"TOTALLYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6

"DOSILION. Cys or Thi in Pl and Ser or Ala in Pl".

"TOTAL CATALLYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                     K
                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N):
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID SET OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                              X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

WEDLINE-9701508B; Pubmed-8861916;

LOVE R.A., Parge H.E., Wickerskam J.A., Hostomsky Z., Habuka N.,

Moomaw E.W., Adachi T., Hostomska Z.;

"The crystal structure of hepatitis C virus NS3 proteinase reveals

trypsin-like fold and a structural zinc binding site.";

Cell 87:331-342(1996).
 phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pro; IPR007095; RNA_pol_DS_PS.
Pro; IPR007094; RNA_pol_PSvir.
PP01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A38465; GNWVTC,
1A1Q; 25-MAR-98,
1JXP; 14-JAN-98,
1NS3; 08-APR-98,
1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1GX6; 10-APR-02.
1CUV; 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1csj;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB;
```

HCV\_NS4a;

PF01006;

Pfam;

```
CELLULAR ANINOPEPPIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTENSE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENT
                                                                                                                                                                                                                                               POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC,
N-LINKED GLCNAC,
N-LIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SW00487; DEXDC; 1.
                                                                                                                                     1103
                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TURN
STRAND
HELIX
TURN
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
                                                                                                                                                                                                                        CHAIN
                                                                                                                                     CHAIN
                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HELIX
TURN
TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24, Last sequence update)
01-AUG-1992 (Rel. 41, Last annocation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP88) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3 4.21.98); Nonstructural protein NSA (P4); Nonstructural protein NSAB (P27); Nonstructural protein NSSA (P56); Nonstructural protein NSSA (P56); CAPSA (PA); Nonstructural protein NSSA (P65); NONSTRUCTURAL NSSA (P65); NONSTRUCTURA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (RNA)(N).

SOBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Suyimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                Length 3010;
                                                                                                                  327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                         Score 144; DB 1; L
Pred. No. 4.7e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                           1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                              95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                            Ouery Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
  1188
1197
1202
1204
1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISCUSSION OF SEQUENCE
1187 118
1189 119
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                         HCVJA
                                                                    TURN
STRAND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HC1
P26662;
  TURN
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                           RESULT B
POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                         EEEEES
                                                                                                                                                                                                                                                        ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions along as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                        01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein [Contains Capsid protein C (Core protein) (P22);
670-68) (GP70) (NSI); Protein F1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NSI); Protein F7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NSA (F49); Nonstructural protein NSAB (P27); Nonstructural P27, NONSTRUCTURAL NSAB (P27); NONSTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDILINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.; "Structure of the hepatitis C virus RNA helicase domain."; Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBDITE: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C. E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS? PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS? PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Prince A.M.; "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates.": Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657
MEDLINE-97331322; Pubmed-9187654;
                   Structure 6:89-100(1998)
                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                               РОГС_НСУН
                                                                                                                                         RESULT 9
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEINS_HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Problem, 12.20.20.

PROLYPTOSELO.

POLYPTOSELO.

POLYPTOSELO.

POLYPTOSELO.

Core protein; Glycoprotein; Transferase; RNA-directed RNA polymerase;

Core protein; Goat protein; Envelope protein; Helicase; ATP-binding;

Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET

I REMOVED FROM CAPSID PROFIEIN C BY THE

CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                            Interpro; IPR007095; RNA_pol_DS_ps.
Interpro; IPR007094; RNA_pol_BSvir.
Pfam; PF01543; HCV_caps1d; 1.
Pfam; PF01539; HCV_cenv; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                                    | IPR002519; HCV_env.
| IPR002531; HCV_NS1.
| IPR002518; HCV_NS2.
| IPR004109; HCV_NS3.
| IPR001495; HCV_NS4a.
| IPR001496; HCV_NS4a.
| IPR001689; HCV_NS5a.
| IPR002868; HCV_NS5a.
| IPR002166; HCV_RGRP.
                                                                                           HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201.
2240
2788
327017 F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00271; hellcase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF02907; HCV_NS3; 1,
PF01006; HCV_NS4a; 1
PF01001; HCV_NS4b; 1
PF01506; HCV_NS5a; 1
                                                                   InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥
HSSP; P26663; 1JXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2014
347
347
1083
1107
11165
11230
1230
234
234
417
423
430
                                                                                                                 InterPro;
InterPro;
                                                                                                                                                                                                                             InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
```

ö

Score 144; DB 1; Length 3010; Pred. No. 4.7e-13;

95.4%; 96.48;

0; Mismatches

Conservative

Best\_Local Similarity Matches 27; Conserv

Query Match

υ

us-09-491-146a-23.rsp

```
Gaps
    ..) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 144; DB 1;
Pred. No. 4.7e-13;
0; Mismatches 1;
  N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3033 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95.4%;
nilarity 96.4%;
Conservative 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
   1636
1652
                                                                                                                                                                                                                                                                                       1377
1380
1385
1393
1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJ6
ID POLG_HCVJ6
  CARBOHYD
CARBOHYD
STRAND
                                  Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
   g
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                            Probon, PD186062, HCC_NS1; 1.
SWART; SMO0487; DEXDC: 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E1.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NSTA.
NONSTRUCTURAL PROTEIN NSTA.
NONSTRUCTURAL PROTEIN NS5A.
NS5A.
NS5A.
NS5A.
NSSA.
NONSTRUCTURAL PROTEIN NS5A.
NSSA.
NSSA.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NSSA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LINKED
LINKED
LINKED
                                                                                                     TRANSFAC, 104155; -
InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR002513; HCV_capaid.
InterPro; IPR002513; HCV_capaid.
InterPro; IPR002513; HCV_capaid.
InterPro; IPR002513; HCV_NS1.
InterPro; IPR00218; HCV_NS3.
InterPro; IPR001049; HCV_NS3.
InterPro; IPR001049; HCV_NS4.
InterPro; IPR001049; HCV_NS4.
InterPro; IPR001049; HCV_NS4.
InterPro; IPR001050; HCV_NS4.
InterPro; IPR001050; HCV_NS4.
InterPro; IPR001050; HCV_NS3.
InterPro; IPR001050; HCV_NS3.
InterPro; IPR00104; RNA_pol_DS_PS.
InterPro; IPR00104; HCV_Capaid; I.
Pfam; PF01560; HCV_NS3; I.
Pfam; PF01560; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF010050; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
PFADD_MS_PF01001; HCV_NS4; I.
PFADD_MS_PF01001; HCV_NS4; I.
                              EMBL; M67463; AAA45534.1;
PIR; A36814; GNWVCH.
                                                    1HEI; 25-NOV-98.
1A1V; 16-FEB-99.
1A1R; 17-JUN-98.
                                                                                    MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
```

Fri Aug

```
1111
1169
1234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SO THE FERRE FOR THE F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92044440; PubMed-1658196;

MEDLINE-92044440; PubMed-1658196;

Machida A., Miyakawa Y., Mayumi M.;

Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.;

J. Gen. Virol. 72:2697-2704(1991).

I. Gen. Virol. 72:2697-2704(1991).

I. Gen. Virol. 72:2697-2704(1991).

C.- FUNATION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C.- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

C.-- CATALITIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                    Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP06) (GR01); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.-)88); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66); Nonstructural protein NS5A (P60); Nonstructural p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS'A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002868; HCV_NS5a.
InterPro; IPR001666; HCV_RGARP.
InterPro; IPR0010560; Halicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF011643; HCV_capaid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR0004109; HCV_NS2.
InterPro; IPR000745; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D00944; BAA00792.1; -. PIR; J01303; J01303. HSSP; P27958; HEI. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4a;
HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01539;
PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01538;
PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 4), Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7) Nonstructural protein NS2 (P21)
(EC 3 4.21.98); Nonstructural protein NS48 (P4): Nonstructural protein
NS48 (P27); Nonstructural protein NS54 (P5); Nonstructural protein
NS58 (P66) (P70) (RNA directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                    CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                           Probon; reactions: Transferase; RNA-directed RNA polymerase; Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBL_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ( POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                            (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. 4.7e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 144; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNTIRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; P186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                        ¥
```

```
MEDLINE-92230206; PubMed-1314449;
MEDLINE-92230206; PubMed-1314449;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termin of viral genomic and antigenomic RNA.";
III The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termin of viral genomic and antigenomic RNA.";
III TOROPHOSIC SUGGESTING A PROFIEND NS2A, NS2B, NS4A AND NS4B ANE HYDROPHOBIC, SUGGESTING A POSSTBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein [Contains: Capsid protein C (Core protein) (P22);
67-68) (GP70) (NS1); Protein F1 (GP32) (GP35); Envelope glycoprotein E2 (GP83) (GP35); Envelope glycoprotein E2 (GP83) (GP35); Envelope glycoprotein E2 (GP83) (GP35); Nonstructural protein NS4 (P4); Nonstructural protein NS4 (P4); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P65); Nonstruct
                      RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                              SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                             CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MW; 1A173E7E3381FD1A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 144; DB 1;
Pred. No. 4.7e-13;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC.
                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                      N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                            ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 90....
-hos 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   649
1091
2038
2359
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
POLG_HCVTW
ID POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                   TRANSMEM
                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P29846;
                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLYPICATION (1) COAT protein; Transferase; RNA-directed RNA polymerase; W Polyprotein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

TREMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIOASE.

TCHAIN 115 CARSID PROTEIN (POTENTIAL).

MATRIX PROTEIN (POTENTIAL).

MATRIX PROTEIN NS1 (POTENTIAL).

TCHAIN 734 1010 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).

TCHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).

TCHAIN 1620 1867 2017 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
                                                                                                                                                                            genotypes."
Virology 188:331-341(1992).

- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

- CATALYTIT CACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

- CATIVITY: A nucleoside triphosphate - N diphosphate +
SEQUENCE FROM N.A.
MEDILINE-92230212; PubMed-1314459;
Okando D., Kural K., Okada S. I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Interpro; IPR002868; HCV_NSSa.
Interpro; IPR007095; RNA_DOL_DS_PS.
Interpro; IPR007095; RNA_DOL_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_caps1d; I.
Pfam; PF01543; HCV_core; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD,
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJB. HSSP; P27598; 1HEI. MERCPS; S29.001; -. MERCPS; 039.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR004109; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <del>:</del>
```

ö

Gaps

Ø

(POTENTIAL). (POTENTIAL). (POTENTIAL).

N-LINKED N-L

(POTENTIAL) ( POTENTIAL (POTENTIAL) (POTENTIAL)

(GLCNAC. (GLCNAC. GLCNAC.

(GLCNAC

N-LINKED (GLCNAC. .) (POTENTIAL)
MW; AAD267D55CDFE215 CRC64;

POTENTIAL)

(GLCNAC (GLCNAC

(POTENTIAL) POTENTIAL (POTENTIAL)

```
623
645
2041
2240
2529
2788
4, 327047
                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                              Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11104;
                                                                  645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                       POLG_HCV1
P26664;
                                      CARBOHYD
CARBOHYD
                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                        CARBOHYD
CARBOHYD
SEQUENCE
           CARBOHYD
CARBOHYD
                                                          CARBOHYD
                                                                                               CARBOHYD
                                                                                                                                                                                                                                     RESULT 13
                                                                                                                                                                                                                                               POLG_HCV1
 g
                                                                                                                                                                                                                                                        ò
                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics. Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CELLULAR ANINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELORE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                 InterPro; IRR002522; HCV_capsid.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002519; HCV_capsid.
InterPro; IPR002519; HCV_capsid.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR0004109; HCV_NS2.
InterPro; IPR000440; HCV_NS4a.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR00166; HCV_NS4b.
InterPro; IPR001096; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01550; HCV_NS1; 1.
Pfam; PF01006; HCV_NS1; 1.
Pfam; PF01006; HCV_NS2; 1.
Pfam; PF01006; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINKED
                                                                                                                 EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                         InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART; SM00487; DEXDC;
                                                                                                                                  PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
                                                                                                                          PIR; A40244; GNWVTW
                                                                                                                                                      MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 347
11003
11107
11165
1230
1230
1316
1316
209
233
305
417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                3D-structure INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-One polyprotein (Contains: Capsid protein C (Core protein) (P22);
67-69 (GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
67-69 (GP70) (NS1); Protein P7; Nonstructural protein NS4 (EC 3.4.22-); Proteases/Helicase NS3 (P70) (Hapacivirin)
62 3.4.21-98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); NONSTRUCTURAL (P65); NONST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
     Length 3010;
                                                                                                         1; Indels
90.7%; Score 137; DB 1; 96.3%; Pred. No. 5.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3011 AA.
                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                         2 KPORKTKRNTIRRPODVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M62321; AAA45676.1; -.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-88288341; PubMed-2840577;
Perrine K.G., Denker J.A., Nilsen T.W.;
"A multi-copy gene encodes a potentially protective antigen in Brugia
                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nilsen T.W., Maroney P.A., Goodwin R.G., Perrine K.G., Denker J.A., Nadduri J., Kazura J.W.;
"Cloning and characterization of a potentially protective antigen in Proc. Natl. Acad. Sci. U.S.A. 85:3604-3607(1988).
                                                                                                                                                                                                                                                                                                                                               PSYN_BRUMA STANDARD; PRT; 548 AA.
P10723;
01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
Asparaginyl-tRNA synthetase, cytoplasmic (EC 6.1.1.22) (Asparagine-tRNA ilgase) (AsnRS) (Potentially protective 63 kDa antigen).
Brugia malayi (Filarial nematode worm).
Busaryota; Mctazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
NCBI_TAXID-6279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-96049569; PubMed-7589498;
Kron M., Marquard K., Hartlein M., Price S., Leberman R.;
"An immunodominant antigen of Brugia malayi is an asparaginyl-tRNA
                                                                                                                  ö
                                                          Length 3011;
        327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                          Score 136; DB 1;
Pred. No. 7.4e-12;
                                                                                                               1; Mismatches
                                                                                                                                                               1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                Mol. Biochem. Parasitol. 30:97-104(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR004522; AsnS.
InterPro; IPR004362; RNA-synt_2.
InterPro; IPR002312; tRNA-synt_asp.
InterPro; IPR004365; tRNA_ant1.
InterPro; IPR006195; tRNA_ligase_II.
Pfam; PF00152; tRNA_Synt_2; 1.
Pfam; PF01336; tRNA_ant1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; J03971; AAA27852.1; -.
EMBL; J03266; AAA27849.1; -.
PIR; A28209; A28209.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-88217950; PubMed-3368467;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRNASYNTHASP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TIGR00457; asnS; 1.
                                                       Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
        3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRINTS; PR01042;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FIGREAMS;
        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FUNCTION.
                                                             Query Match
                                                                                                                                                                    à
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                          HE TO THE TRANSPORT OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJOR ENFOREIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2/E2 (POTENTIAL).
PROTENSE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY S:
CHARGE RELAY SYSTEM (BY S)
CHARGE RELAY SYSTEM (BY S)
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC, N-LINKED) (GLCNAC, N-LINKED (GLCNAC, N-LINKED) (GLCNAC, N-LI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                              InterPro; IFR00252; HCV_capsid.
InterPro; IFR00252; HCV_capsid.
InterPro; IFR00252; HCV_capsid.
InterPro; IFR00251; HCV_capsid.
InterPro; IFR00251; HCV_NS1.
InterPro; IFR002518; HCV_NS2.
InterPro; IFR001490; HCV_NS3.
InterPro; IFR001490; HCV_NS3.
InterPro; IFR001490; HCV_NS4.
InterPro; IFR001490; HCV_NS5a.
InterPro; IFR001650; Helicase_C.
InterPro; IFR001650; Helicase_C.
InterPro; IFR001650; Helicase_C.
InterPro; IFR001650; Helicase_C.
InterPro; IFR001650; HCV_Capsid; I.
Pfam; PF01543; HCV_Capsid; I.
Pfam; PF01543; HCV_Capsid; I.
Pfam; PF01540; HCV_NS2; I.
Pfam; PF01006; HCV_NS2; I.
Pfam; PF01001; HCV_NS2; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS5a; I.
Pfam; PF00101; HCV_NS5a; I.
Pfam; PF00101; HCV_NS5a; I.
Pfam; PF00101; HCV_NS5a; I.
                                                                                                                             InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
InterPro; IPR002521; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00998; Viral_R
Probom; PD186062; HCV_
SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                645
2041
2077
2240
2364
2789
                         1A1V; 16-FEB-99
1HEI; 25-NOV-98
A39166; GNWVC3
                                                                           MEROPS: S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     347
1083
11165
11165
11230
1230
209
234
305
417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Leduc R., Molloy S.S., Thorne B.A., Thomas G.;
"Activation of human furin precursor processing endoprotease occurs by
an intramolecular autoproteolytic cleavage.";
J. Biol. Chem. 267:14304-14308(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mode1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Brake A.J.;
CDNA and gene structure for a human subtilisin-like protease with
cleavage specificity for paired basic amino acid residues.";
DNA Cell Biol. 10:319-328(1991).
                                                                                                                                                                                                                                                                                                                                                FURI_HUMAN STANDARD; PRT; 794 AA.
P09958; 014336;
01-AR-1989 (Rel. 10, Created)
01-AR-1990 (Rel. 14, Last sequence update)
01-AR-1990 (Rel. 42, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue cleaving enzyme) (PACE) (Dibasic processing enzyme).
FURIN OR PACE OR FUR.
FURIN OR PACE OR FUR.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                 IMMUNOLOGICAL REACTIVITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D., Dorssers L.C.J., van de Ven W.J.M.;
"Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2.";
Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.;
Evolutionary conserved close linkage of the c-fes/fps proto-o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-94291619; PubMed-8020465; Siezen R.J., Creemers J.W.M., van de Ven W.J.M.; Incompose modelling of the catalytic domain of human furin. If for the eukaryotic subtilisin-like proprotein convertases."; Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                            DB 1; Length 548;
                                            REGION OF IMMUNOLOGICAL REAC

I -> V (IN REF. 2).

H -> S (IN REF. 2).

L -> V (IN REF. 2).

L -> V (IN REF. 2).

Q -> K (IN REF. 2).

Q -> K (IN REF. 2).
                                                                                                                                                                                                           10; Indels
                                                                                                                                                                                                         4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -STRUCTURE MODELING OF CATALYTIC DOMAIN.
                                                                                                                                                                          Score 54;
Pred, No.
                                                                                                                                                                                                                                                           4 QRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                            Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 296-794 FROM N.A. MEDLINE-87053858; Pubmed-3023061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-91321735; PubMed-1713771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92332543; PubMed-162922;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-90175002; PubMed-2408021;
                                              131 RE
50 I
65 H
212 P
265 L
454 O
62339 MW;
                                                                                                                                                                        sh 35.8%;
| Similarity 44.0%;
| 11; Conservative
                                            81
50
65
212
265
454
454 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE-Blood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROCESSING
                                              DOMAÍN
CONFLICT
CONFLICT
                                                                                                         CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                       450
                                                                                                                                                                            Query Match
                             Antigen.
                                                                                              CONFLICT
                                                                                                                                                                                           Local
                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                     ULT
                                                                                                                                                                                                                                                                   8
 NA KEN
KEN
FT FT FT SO
                                                                                                                                                                                                                                         ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics. Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE CACTUATY WITHIN CONSTITUTIVE SECRETORY PARHWAYS AND CAPABLE OF CLEAVAGE AT THE RX(K/K)R CONSENSUS MOTIF.
CATALYTIC ACTIVITY: Release of mature proteins from their proproteins by cleavage of Arg-Yaa-Yaa-Arg-1-Zaa bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and von Willebrand factor from their
                                                                                                                                                                                          COFACTOR: CALCIUM-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Transmembrane; Glycoprotein; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY SCHARGE RELAY SYSTEM (BY SCHARGE RELAY SYSTEM (BY SPOTENTIAL.)

POTENTIAL.

POTENTIAL.

N-LINKED (GLCNAC. ..) (FOR THE CONTROL OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GO; GO:0005794; C:Golgi apparatus; TAS.
GO; GO:0004276; F:furin activity; TAS.
GO; GO:0007267; P:cell-cell signaling; TAS.
GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CYS-RICH.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF01483; P_proprotein; PARTIAL.
Pfam; PF00082; Peptidase_58; 1.
PRINTS; PR00723; SUBTILISIN.
ProDom; PD000717; P_domain; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00136; SUBTILASE_ASP; 1.
PROSITE; PS00137; SUBTILASE_HIS; 1.
PROSITE; PS00138; SUBTILASE_SER; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR006212; Furin_repeat.
InterPro; IPR002884; P_domain.
InterPro; IPR000209; Peptidase_S8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -. PIR; A39552; KXHUF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protease;
                                                                                                                                                                      respective precursors.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FU;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1MPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hydrolase; Serine
Zymogen; Calcium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S08.071;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00261;
                                                                                                                                                                                                                                         PROPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            099405;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IGNAL
```

```
ö
                                                                                                                                              Gaps
                                                                                                                                              ;
0
70 75 CLEAVAGE (SECOND AUTO-).
104 107 CLEAVAGE (FIRST AUTO-).
498 500 CELL ATTACHMENT SITE (POTENTIAL).
759 762 CELL SURFACE SIGNAL.
773 779 TRANS GOLGI NETWORK SIGNAL.
774 AA: 86678 MW; 10C44DD5892EF85D CRC64;
                                                                                                           Ouery Match 34.4%; Score 52; DB 1; Length 794; Best Local Similarity 50.0%; Pred. No. 6.8; Matches 9; Conservative 4; Mismatches 5; Indels
 SITE
SITE
SITE
SITE
SITE
SEQUENCE
 FT
FT
FT
SO
```

4 ORKTKRNTIRRPODVKFP 21 :|:|||: | | ||| 102 KRRTKRDVYGEPTDPKFP 119 ò

g

Search completed: August 7, 2003, 11:19:55 Job time : 6.90909 secs

```
Eri Aug 8 09:20:02 2003
```

									8 00 00 00 00 00 00 00 00 00 00 00 00 00	R R R	10008 10008	SRREEN Y G
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.	protein - protein search, using sw model	<pre>Run on: August 7, 2003, 11:05:41 ; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec</pre>	Title: US-09-491-146A-23 Perfect score: 151 Sequence: 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28	Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5	Searched: 830525 seqs, 258052604 residues	al number of hits satisfying chosen parameters: 830525	Minimum DB seq length: 0 Maximum DB seq length: 2000000000	Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries	Database : SPTREMBL_23:*  1: sp_archea:* 2: sp_bacteria:* 3: sp_fungi:* 4: sp_human:* 5: sp_human:* 6: sp_nammal:* 7: sp_mncalle:* 8: sp_organelle:* 9: sp_phage:* 10: sp_phage:* 11: sp_rodent:* 12: sp_virus:* 13: sp_vertebrate:* 14: sp_unclassified:* 15: sp_virus:* 16: sp_bacteriap:* 17: sp_archeap:*	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	SUMMARIES Result Query No. Score Match Length DB ID	1         151         100.0         45         12         Q68307         Departitis           2         151         100.0         45         12         Q68310         Q68310         Departitis           3         151         100.0         45         12         Q68310         Q68310         Departitis           4         151         100.0         61         12         Q81782         Q81782         Q81782         Departitis           5         151         100.0         61         12         Q81784         Departitis         Q81786         Departitis         Q81877         Departitis         Q81786         Departitis </td

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		0;
hepatiti, hepatiti, hepatiti	dae; dae; ltis C D BY A SS:	Gaps
0899981 0899983 0899983 089334 089334 089334 089600 08900 08900 089000 0	Flaviviridae Sof hepatitis Ses. WO PROTEINS: S A COMPLEX OF	4; Length 45; Indels 0;
		12; ; 0;
	ALIGNMENTS  ated) t sequence upd t annotation up n) (Fragment). viruses, no DN tuses, no DN VIRUS IS A NU NVELOPE CONSIS E. THE NUCLEOC LARITY).	151; 1 No. 1e natches 7G 28
080P81 080P83 080P76 080P76 080M14 068893 080ECX1 08BECX9 08BECW9 08BECW9 08BECW9 08BECW9 08BECW9 08BECW9 08BECW9 08BECW9 08BECW9 08BECW1 08BE	ALIGNMENT  ALIGNMENT  INT. PRT; 45  el. 01, Created)  el. 22, Last sequencel. 22, Last sequencel. 22, Last annotation of the strand viruses,  live-strand viruses,  to the EMBL/GenBank  to the EMBL/GenBank  to the EMBL/GenBank  THE ENVELOPE  THE ENVELOPE  THE STAIL ANTERNATIT).  152.1; 75.21, 17.006  164. Mai. CC522167006  164. Mai. CC522167006  165.1; 74. Mai. CC522167006  165.1; 74. Mai. CC522167006  166.1 Mai. CC522167006  167.1 Mai	43 AA; 3014 NW; CC32/10/090 100.0%; Score 151; imilarity 100.0%; Pred. No. ; Conservative 0; Mismatch PRPORKTRRNTIRRPODVKFPGGGGIVG 28
	), Cr. 01, Cr. 22, La. 22, La. 1yprote -strand -strand Cr. THE EM OF THE PROFEIN (PROFEIN 1;	100.0%; 100.0%; 1ve :rrpodvk 
0000 0000 1000 1000 1000 1000 1000 100	PRELIMINARY;  07.  07-1996 (TrEMBLrel. 01,  07-1996 (TrEMBLrel. 22,  1-1156 (TrEMBLrel. 21,  1-1156 (MAR-1995) to the  1-156 (MAR-1995) to the  1-157 (MAR-1995) to the  1-156 (MAR-1995) to the  1-157 (MAR-	100 100 100 100 100 100 100 100 100 100
000000000000000000000000000000000000000	RELIMIN TTEMBLY TTEMBLY TTEMBLY (Genome irus. A posit N.A. 21; N.A. 21; N.A. AND ME AND ME AN	Darity Conser QRKTKR
	PRELIM PRELIM PRELIM 96 (TrEMB 00 (TrEMB 01 (Geno 12 (Viss.) 12 (Viss.) 13 (Viss.) 14 (MAR-199 111 (MAR-199 111 M M AND 111 M M AND 111 M M AND 111 M AND 11 M AND	Similarity Similarity 8, Conser PRPORKTRR
	ULT 1  068307 068307 068307 01-NOV-1996 (TrEMBI 01-NOV-1996 (TrEMBI 01-OCT-2002 (TREMB	atch sals sals 1
1111000000000000000000000000000000000	888	Query Mo Query Mo Best Loc Matches
	# 9 U A F F F B 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	λο

DR DR DR

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
EMBL, AF506613; AAM33389-1; -
InterPro; IPRO02522; HCV_capsid.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Generic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genenk/DBB Jatebases.
-1-SUBUNIT: THE VIRION OF THIS YIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN MAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
EMBL, AF506612; AAM33381.; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 151; DB 12; Length 60; 100.0%; Pred. No. 1.4e-15; Live 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60 AA; 6725 MW; ACCE7D9C8B90299A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B92CCE7D9C8B9029 CRC64;
                                                                                                                                                                                                                                                                                                                                            (TrEMBLrel. 23, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 AA.
                                                                                                                                                                                                                                                          60 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORE PROTEIN.
                                                  5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                 1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last seq
01-MAR-2003 (TrEMBLrel. 23, Last ann
                                                                                                                                                                                                                                                                                                                          (TrEMBLrel. 22, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
6881 MW;
                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2003 (TrEMBLrel, 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. STRAIN-RIG256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-RIG253;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
CHAIN
                                                                                                                                                                                                                                                                                                                   01-0CT-2002
01-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                   08JYS2
08JYS2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8JYS3
                                                                                                                                                                                  RESULT 4
Q8JYS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QBJYS3
                                                                                                                                                                                                                                                                                       SPERMENT STATE OF STATE OF SPERMENT OF SPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Songsivilai S., Kanistanon D., Kunkitti R., "Identification and characterisation of Thai isolates of hepatitis {\tt C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LICOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
--- PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF EMBL; U33746; AA65053.1;
--- InterPro; IPR002522; HCV_capsid.
--- Prome Profession of the Profession
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
10-OCT-2011 (Genome polyprotein) (Fragment).
10-OCT-2012 (Vills).
11-OCT-2013 STRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINK [BY SIMILARITY).

EMBL; U23748; AAA65055.1; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Score 151; DB 12; Length 45; 100.0%; Pred. No. 1e-15; Live 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 12; Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER 45 45
SEQUENCE 45 AA; 5014 MW; CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     45 45 45
45 AA; 5014 MW; CC527167096DDAF6 CRC64;
                                                                                                           ol-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Score 151; DB 12
100.0%; Pred. No. 1e-15;
tive 0; Mismatches 0
                                                                              45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 100.
es 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-HCV-C94009;
                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HCV-BB38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                       968308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   068310
                                                                    908890
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            virus.
RESULT 2
Q68308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   068310
```

셤

ç

OPPTE

ö

Gaps

ö

Ø

K

Gaps

g

ò 8 

```
STRAIN-233;
MEDLINE-21904745; PubMed-11907242;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
J. Virol. 76:4634-4043(2002).
I. VIROL. 76:4634-4043(2002).
PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AV070178; AAL58592.1;
InterPro: IPR002522; HCV_capsid.
Pfam; PP01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=609;
MEDLINE-21904745; Pubmed-11907242;
Kallinne O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 100;
                                                                                                                     Length 100;
                                                                                                                                                              Indels
                                                               100 100
100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100
11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                     01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 151; DB 12;
100.0%; Pred. No. 2.5e-15;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
                                                                                                                   100.0%; Score 151; DB 12;
100.0%; Pred. No. 2.5e-15;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                            100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100 AA.
                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last seq
01-OCT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
  InterPro; IPR002522; HCV_caps1d.
Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.08;
                                                                                                                                         Local Similarity 100.0
es 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
nes 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                         100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100
                                    Polyprotein.
NON_TER 1(
SEQUENCE 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER 10
SEQUENCE 100
                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                               080P87
                                                                                                                                                                                                                                                                                                                                          Q8QP87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            080P72
                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     σ
                                                                                                                                                                                                                                                                                                 RESULT 6
Q8QP87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NA WAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
-1. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN BNVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AY070180; AAL58594.1; -...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DBBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROFIIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL, AF506621; ARAM33997.1; -
InterPro: IPR002522; HCV_capsid.
Prim: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Score 151; DB 12; Length 73; 100.0%; Pred. No. 1.8e-15; ive 0; Mismatches 0; Indels (
    Length 61;
                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6EDC082DAE8CB796 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
  100.0%; Score 151; DB 12;
100.0%; Pred. No. 1.4e-15;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-2002 (TrEMBLrel. 21, Last sequence upda
01-0CT-2002 (TrEMBLrel. 22, Last annotation update protein (Genome polyprotein) (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 AA.
                                                                                                                                                                                                                       73 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                    01-OCT-2002 (TrEMBLrel. 22, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-602;
MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                      Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      73 73
73 AA; 8201 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-2002 (TrEMBLrel. 21,
01-JUN-2002 (TrEMBLrel. 21,
01-OCT-2002 (TrEMBLrel. 22,
Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 100.
Les 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. STRAIN-RIG291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                               01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
                                                                                                                                                                                                                     QBJYR4
QBJYR4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8QP85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                            RESULT 6
QBJYR4
```

```
100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100 AA;
                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                    100
                                                                                                                                                                                                                                                                                      Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8QP77
Q8QP77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8QP80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q8QP80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 080P80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          080P77
            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RA MEDLINE-190475; Pubmed-11907242;

RA Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

RA Andrural intergency procedulant of hepatitis C virus identified

RT in St. Petersburge."; Carter and C
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                           100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; Live 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER 100 100
SEQUENCE 100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                          SEQUENCE 100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 151; DB 12;
Pred. No. 2.5e-15;
Whishatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                       100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100 AA
                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.08;
                                                                                                                                                                                                                                                               28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                            100
                                                                                                                                                            Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                    Q8QP71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8QP74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11

080P74

ID 080P7

AC 080P7

DT 01-JU

DT 01-JU

DT 01-OC

DE CORE

OS HEPAL
                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                   SULT 10
     ò
                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                      S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
```

```
C STRAIN-631;

XX MEDLINE-21904745; PubMed-11907242;

XA Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

Ra Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

Ra natural intergency T. recombinant of hepatitis C virus identified in St. Petersburg.";

YICL O. 76:4034-4043(2002).

C ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA BY SIMILARITY).

CC PROTEIN C AND MRNA BY SIMILARITY).

DR EMBL; ARVO0191; AALS8605.1; --

DR EMBL; ARVO1091; AALS8605.1; --

DR FMDL; PRO1043; HCV_capsid.

DR PAM: PF0143; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-21904745; PubMed-11907242;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergencypic recombinant of hepatitis C virus identified
in St. Petersburg ";
J. Virol. 76:4034-4043(2002).
-!- SUBUNIT: THE VIRON OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
IIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AY070185; AALS8999.1;
-INTERPRO: IPR002222; HCV_capsid.
PP01543; HCV_capsid, 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; ive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 151; DB 12; Length.100; 100.0%; Pred. No. 2.5e-15; 1.ve 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                      100 | 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100
11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence upds
01-JUN-2002 (TrEMBLrel. 22, Last annotation up
01-OCT-Drotein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPORKTKRNTIRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTIRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
```

```
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

"A natural intergenctypic recombinant of hepatitis C virus identified
In St. Petersburg;
J. Virol. 76:4034-4043(2002).

-1. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AYO70190; AAL58604.1;
R InterPro; IPR002522; HCV.capsid.
R InterPro; IPR002522; HCV.capsid.
W PFIM: PFO1543; HCV.capsid; 1.
W POLYPOTOTEN

NON.TER
OND.TER
OND.TER
SEQUENCE 100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                            01-JUN-2002 (TrEMBLrel. 21, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 151; DB 12;
100.0%; Pred. No. 2.5e-15;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                               Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
     5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Search completed: August 7, 2003, 11:18:55 Job time: 28.6364 secs
                                                                                                                                                                                   01-JUN-2002 (TrEMBLrel. 21, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-680;
MEDLINE-21904745; Pubmed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                   080P75
                                                                                   RESULT 15
Q8QP75
                                                                                                                                                              STATE TO CO CONTRACTOR OF THE PROPERTY OF THE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FURINE=118;
MEDLINE=21904745; PubMed=11907242;
MEDLINE=21904745; PubMed=11907242;
MEDLINE=21904745; PubMed=11907242;
Malurina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenortypic recombinant of hepatitis C virus identified in St. Petersburg.";
J. Virlol. 76:4034-4043(2002)
-1. SIGNATT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SMILLARITY).
EMBL; AY070181; AAL58595.1; -.
                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Hepatitis (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; ive 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 151; DB 12; Length 100; 100.0%; Pred. No. 2.5e-15; tive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100 100
100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 100.
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8QP84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
Q8QP84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CSSSEPPERSS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STED
```

Gaps

.

1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28

ò

ô

ö

Length 100; Indels

```
PCT-US95-13660-3
 61
95
144
28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 187, App
Sequence 188, App
Sequence 188, App
Sequence 189, App
Sequence 190, App
Sequence 230, App
Sequence 189, App
Sequence 189, App
Sequence 189, App
Sequence 175, App
Sequence 175, App
Sequence 175, App
Sequence 175, App
Sequence 24, Appli
Sequence 26, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 28, Appli
Sequence 27, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 23, Appl
Sequence 187, App
                                                                                          August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                         Issued_Patents_AA:*
.: /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PcTUS_COMB.pep:*
.: /cgn2_6/ptodata/1/laa/PcTUS_COMB.pep:*
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-921-887-23
US-08-290-665A-188
US-08-290-665A-188
US-08-290-665A-188
PCT-US95-10398-188
PCT-US95-10398-188
US-08-836-075A-44
US-08-836-075A-44
US-08-290-665A-189
PCT-US95-10398-189
PCT-US95-10398-189
US-08-290-665A-175
PCT-US95-10398-175
US-08-280-665A-175
US-08-280-665A-175
US-08-380-160-2
US-08-380-160-2
US-08-380-160-2
US-08-380-160-1
                                                                                                                                                                                                                                                                            al number of hits satisfying chosen parameters:
                                                                                                                                                                             1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                    328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                    protein search, using sw model
                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                             US-09-491-146A-23
151
                                                                                                                                                                                                                                                                                                             seq length: 0 seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           %
Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0
                                                                                                                                                                Perfect score:
Sequence:
                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                          Minimum DB :
Maximum DB :
                                                                   OM protein
                                                                                                                                                                                                                                                                                                                                                                                                             Database :
                                                                                                                                                                                                                                                     Searched:
                                                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Title:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
```

PCT-US93-08638-9 PCT-US94-07088-23

Sequence 3, Appli Sequence 10, Appl Sequence 132, App Sequence 14, Appl Sequence 8, Appli Sequence 8, Appli Sequence 8, Appli Sequence 12, Appli Sequence 15, Appli Sequence 16, Appli Sequence 16, Appli Sequence 16, Appli Sequence 17, Appli Sequence 16, Appli Sequence 17, Appli Sequence 16, Appli Sequence 17, Appli Sequence 17, Appli Sequence 17, Appli Sequence 17, Appli	FOR MAKING THE SAL	0; Gaps 0;
	CTION METHOD OOF	; Length 28 5; 0; Indels
PCT-US95-13660-3 US-08-635-866-198 US-08-635-8666-198 US-08-635-8666-232 US-08-34-977-8 US-08-34-977-8 US-09-904-686A-8 US-09-904-686A-8 US-09-501-195-2 US-08-501-195-2 US-08-816-055-46 US-08-07-05-135-5-2	ALIGNMENTS  1 US/08921887  1, YURI E. 190WARD A. 190WARD A. 190WARD A. 190MARD A. 190MARD A. 190MARD A. 190MARD A. 190MOSIC PROTEIN AND RESTRICTION 190MOSIC PROTEIN AND RESTRICTION 190MOSIC PROTEIN AND RESTRICTION 190MOSIC PROTEIN AND RESTRICTION 190MOSIC AND	Score 151; DB 3; Pred. No. 3.3e-15 Mismatches (
61 5 PCT 74 3 US- 100 4 US- 100 3 US- 1115 1 US- 1115 2 US- 1124 1 US- 1137 3 US- 1137 3 US- 1140 2 US- 1154 3 US- 1156 3	ALIG  "-91-887-23  "-921-887-23  "Hence 23, Application US/08921887  Ent No. 6030771  "NERAL INFORMATION: APPLICANT: FIELDS, HOWARD A.  TITLE OF INVENTION: MOSAIC PROTEIN  NUMBER OF SEQUENCES: 55  CORRESPONDERS: 55  CORRESPONDERS: 191 Peachtree Street, N.I  CITY: Allanta  STREET: 191 Peachtree Street, N.I  CITY: Allanta  STREET: 191 Peachtree Street, N.I  COUNTRY: USA  COUNTRY: USA  COUNTRY: IBM PC COMPATA: MEDIUM TYPE: Floppy disk  COMPUTER: IBM PC COMPATA:  MEDIUM TYPE: Ploppy disk  COMPUTER: IBM PC COMPATION:  MEDIUM TYPE: PLOPPY disk  CONFRATION NUMBER: US/08/921,88  FILLNG DATE:  CLASSIFICATION NUMBER: US/08/921,88  FILLNG DATE:  CLASSIFICATION INFORMATION:  TELEPHONE: 404-818-3700  TELEPHONE: 404-818-3700  TELEPHONE: A 404-818-3700  TYPE: A AND OCITE OF SEQUENCE  TYPE: AND OCITE OF SEQUENCE  TYPE: AND OCITE OF SEQUENCE  ORRANISM: HEPATITIS VITUS  "SOCIETING SURCE: OCITE OR SURCE: ORGANISM: HEPATITIS VITUS	. 0 . 8 . 0 . 0
	1-91-887-23 1uence 23, Application US/08921 lent NO. 6030771 INBRAL INFORMATION: APPLICANT: KHUDYAROV, YURI E. APPLICANT: MADRESSE: ADORESSE: 55 CORRESPONDENCE ADDRESS: 55 CORRESPONDENCE S. 55 CORRESPONDENCE S. 55 CORPESSE: 191 Peachtree Stree SITTE: 41 Aanta STREET: 191 Peachtree Stree SITTE: 30303-1769 COUNTRY: GA COUNTRY: USA COUNTRY: USA COUNTRY: BE PROPY GAS COMPUTER: PAPE Floppy disk COMPUTER: PAPE FLOPPY disk COMPUTER: PAPE FLOPPY disk COMPUTER: PAPE FLOPS/MS SOFTWARE: PATENTION DATA: MEDIUM TYPE: PATENTION DATA: FILING DATE: CLASSIFICATION NUMBER: US/08/9 FILING DATE: CLASSIFICATION NUMBER: 36,714 REGISTRATION NUMBER: 36,7	h Similarity 100 28; Conservative
288 299 390 390 390 390 390 490 490 490 490 490 490 490 490 490 4	RESULT 1  US-08-921-887-23  Sequence 23, Application US/08921887  PATENT NO. 6030771  GENERAL INFORMATION: APPLICANT: FIELDS, HOWARD A. TITLE OF INVENTION: ENDONUCLEASY NUMBER OF SEQUENCES: 55 CORRESPONDENCE ADDRESS: CORRESPONDENCE ADDRESS: CORRESPONDENCE ADDRESS: CORRESPONDENCE ADDRESS: COUNTRY: USA ZIP: 30303-1769 COUNTRY: USA ZIP: 30303-1769 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER READABLE FORM: MEDIUM TYPE: Ploppy disk COMPUTER: DAP PC COMPATION OPPERATING SYSTEM: PC-DOS/MS-DS SOFTWARE: PATENTION DATA: APPLICATION NUMBER: US/08/921, FILLNG DATE: CLASSIFICATION NUMBER: US/08/921, FILLNG DATE: CLASSIFICATION NUMBER: 0306: TELEPAX: 404-818-3709 INFORMATION FOR SEQ ID NO: Z3: SEQUENCE CHARACTERISTICS: LENGTH 28 amino acids TYPE: amino acid STRANDEDNESS: NOT relevant MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-SENSE: NO ORIGINAL SURCE: ORGANISM: HEPATILIS VITUS US-08-921-887-23	Query Match Best Local S Matches 28
·	A D C C C C C C C C C C C C C C C C C C	

ò

```
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HERATITIS C VIRUS
NUT THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 151; DB 2;
Pred. No. 2.6e-14;
; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BUKH, J., MILLER, R.H.
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: AND THE USE OF REAGE
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: ADRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIOW TYRE:
COMPUTER:
COMPUTER:
DBM PC COMPATIBLE
OPERATING SYSTEM:
SOFTWARE:
WORDERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER:
US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGRAT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 3026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEPHONE: (212) 758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/290,665A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 190, Application US/08290665A Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IBM PC COMPATIBLE
SYSTEM: PC-DOS/MS-DOS
WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Sci
100.0%; Pri
tive 0;
                                                            ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 421792
TELEX: 421792
INFORMATION FOR SEQ ID NO: 188: SEQUENCE CHARACTERISTICS: LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                 ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 100.0
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S'
                 NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
  TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-290-665A-190
                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 100.0%; Score 151; DB 2; Best Local Similarity 100.0%; Pred. No. 2.6e-14; Matches 28; Conservative 0; Mismatches 0;
                 1 PKPORKTKRNTIRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQKKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECHANICATION INFORMATION:
TELECHAN: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ. ID NO: 187: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRNTIRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ., MILLER, R.H. AND
                                                                                                                                         Sequence 187, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 188, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                    MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HK10
                                                                                                                                                                                                                                                                                                                                                                                                                      : 345 PARK AVENUE
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: BUKH, J.,
APPLICANT: PURCELL, R
TITLE OF INVENTION: N
TITLE OF INVENTION: C
TITLE OF INVENTION: C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-290-665A-187
                                                                                                                    US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 3
US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                               g
```

a

ô

ö

Gaps

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 191;
                                                                                                                                                                                                                                                 Query Match
100.0%; Score 151; DB 5;
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 151;
                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
RPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 188, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AGGUST 1994
ATORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 36,459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.08;
    SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                  ORGANISM: homosaplens
; INDIVIDUAL ISOLATE: HK10
PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INDIVIDUAL ISOLATE: S52
PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: un
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: U:
ZIP: 10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
100.0%; Score 151; DB 2;
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELEPHONE: (212) 758-4800
TELEPHONE: (212) 758-4800
TELEPHONE: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
FAPLICATION UNMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION DATA:
PRIOR APPLICATION DATA:
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 187, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 187:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (212) 758-4800
(212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: homosapiens
INDIVIDUAL ISOLATE: DK12
-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: 345 PARK AVENUE CITY: NEW YORK STATE: NEW YORK COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: unknown ORIGINAL SOURCE: ORGANITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
LIEVEN
NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
APPLICANT: MARRTENS, Geert
TITLE OF INVENTION: IMMUNODMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: UNBER
TITLE OF INVENTION: UNBER
CURRENT PELICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                     CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COUNTRY: USA
ZIP: 77210-4433
COUNTRY: USA
ZIP: 77210-4433
COUNTRY: USA
COUNTRY: ENABLE FORM:
MEDIUM TYPE: Flory disk
COMPUTER: IBM PC compatible
OPERATING SYSTE: C-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA: APPLICATION DATA:
APPLICATION DATA: PCT/EP95/04155
FILING DATE: 23 Oct 1995
FILING DATE: 23 Oct 1995
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
APPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 151;
Pred. No. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPORKTKRNTIRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                             ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 230, Application US/08635886C Patent No. 6555114 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%;
100.0%;
                                                                                                                                                                                           AGENTS
207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 319 amino acids
TYPE: amino acid
                     GENERAL INFORMATION:
APPLICANT: MAERTENS, GEE
APPLICANT: STUVER, LIEV
TITLE OF INVENTION: NEW
TITLE OF INVENTION: AND
TITLE OF INVENTION: AND
TITLE OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 28; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                         AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HERATITIE C VIRUS MAID THE USE OF REAGENYE DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 191;
                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Indels
   100.0%; Pred. No. 2.6e-14;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 151; DB 5;
100.0%; Pred. No. 2.6e-14;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                             NUCLEOTIDE AND DEDUCED
                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                          PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION DATA:
PRICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
FROM APPLICATION DATA: 08/200/665
FILING DATE: 15 AUGUST 1994
APPLICATION NUMBER: 08/200/665
FILING DATE: 15 AUGUST 1994
APPLICATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
FELEVAL: (212) 751-6849
TELERA: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                              APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                 PCT-US95-10398-190
; Sequence 190, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
US-08-836-075A-44
; Sequence 44, Application US/08836075A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: homosapiens
INDIVIDUAL ISOLATE: DK12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : 345 PARK AVENUE
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 191 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 100.0
Best Local Similarity 100.0
Matches 28; Conservative
Best Local Similarity 100.0
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: MUCI
TITLE OF INVENTION: AMIN
TITLE OF INVENTION: CORE
TITLE OF INVENTION: SEQUENCES: SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM: MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

Gaps

ö

: PatentIn version 3.1

FEATURE:

```
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: MUCLECTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 99.3%; Score 150; DB 2; Best Local Similarity 96.4%; Pred. No. 3.6e-14; Matches 27; Conservative 1; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRNTIRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTIRRPQDIKFPGGGIVG 32
                                                                                      SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECHMONICATION INFORMATION:
TELECHMONE: (212) 758-4800
TELEFRAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 189, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                       PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3: MORGAN & FINNEGAN
345 PARK AVENUE
                                              COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                TELEX: 421792
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
                      FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: homosaplens; INDIVIDUAL ISOLATE: S2 US-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: MUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 151; DB 4; 100.0%; Pred No. 4.5e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
LOCATION: (174)..(174)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: MISC_FEATURE
COCATION: (233)
COTHER INFORMATION: Xaa is any amino acid
US-08-635-886C-230
                                                                                                                                                                                                                                                               NAME/KEY: MISC_FEATURE
LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: MISC_FEATURE
LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                    LOCATION: (144)..(144)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                             LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                     COCATION: (161)..(161)

THER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: (167)...(167).
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: (171)..(172)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: (232). (232)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 189, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK
IYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 100.0
Warches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURE:
NAME/KEY: MISC_FEATURE
                                                                         NAME/KEY: MISC_FEATURE
                                                                                                                                                                       NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                  COCATION:
```

FEATURE:

g

us-09-491-146a-23.rai

```
APPLICANT: LERONTS. Geert
APPLICANT: LERONTS. Geert
APPLICANT: LERONTS. Geert
APPLICANT: LERONTS. Geert
APPLICANT: MAERTENS, Geert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNDODMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VINUS
FILE REFERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR PELING DATE: 1993-11-04
NUMBER OF SEO ID NOS: 286
SOFTWARE: PRESENCE: 1933-11-04
SEO ID NO 191
SEO ID NO 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: DELEYS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT FILING DATE: 1996-04-25
PRIOR PELICATION NUMBER: PCT/EP94/03555
PRIOR PLICATION NUMBER: PP 93402718.6
PRIOR PLICH DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version 3.1
SEQ ID NO 192
LENGTH: 450
                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                     Score 146; DB 3; Length 28; Pred. No. 1.7e-14; 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 146; DB 4; Length 45
Pred. No. 3.4e-13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 96.7%; Score 146; DB 4; Best Local Similarity 96.4%; Pred. No. 3.4e-13; Matches 27; Conservative 0; Mismatches 1
                                                                                                                                           1 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 28
                                                                                                                       1 PKPORKTKRNTIRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                   Sequence 191, Application US/08635886C Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 192, Application US/08635886C
Patent No. 6555114
                                   Ouery Match 96.7%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; ORGANISM: hepatitis C virus
US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TYPE: PRT; ORGANISM: hepatitis C virus US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 96.7
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
                                                                                                                                                                                                                                                 US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                         ò
                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 25, Application US/08921887
Fatent No. 6030771
GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                        Score 150; DB 5; Length 191;
Pred. No. 3.6e-14;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEPHONE: (212) 751-6849
TELEPEX: (212) 751-6849
TELEPEX: 421792
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CRARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      36,714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                          99.3%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 25:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Hepatitis virus US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA ZIP: 30303-1769 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,7
                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 96.4*
Thes 27; Conservative
                                                                                                                                                                                                                                                                                                    ORGANISM: homosapiens INDIVIDUAL ISOLATE: S2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ANTI-SENSE: NO FRAGMENT TYPE: i ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Atlanta
                                                                                                                                                                                                                                                                                                                                                PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF ISOLATES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match

96.0%; Score 145; DB 2; Length 191;
Best Local Similarity 96.4%; Pred. No. 1.9e-13;
Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IN PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOSTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026/4116
FILING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026-4116
FILING DATE: 15-ANG-1994
CLASSIFICATION NUMBER: 2026-4116
FILEPHONE: (212) 751-6849
FELEPHONE: (212) 751-6849
FELEPHONE: (212) 751-6849
FELEFAX: (212) 751-6849
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                            Sequence 175, Application US/08290665A; Patent No. 5882852; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: homosapiens;
INDIVIDUAL ISOLATE: P8
US-08-290-665A-175
                                                                                                                                                                                                                                           US-08-290-665A-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
            ô
                                                                              q
```

Search completed: August 7, 2003, 11:23:50 Job time : 11.6364 secs

us-09-491-146a-23.rapb

```
(without alignments)
231.506 Million cell updates/sec
                                                                                                                                                                                    August 7, 2003, 11:20:11; Search time 14.3636 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Publich and Applications_Ark ...

1. (egn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2. (egn2_6/ptodata/1/pubpaa/US06_WEW_PUB.pep:*
3. (egn2_6/ptodata/1/pubpaa/US06_WEW_PUB.pep:*
3. (egn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5. (egn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5. (egn2_6/ptodata/1/pubpaa/US08_WEW_PUB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
7. (egn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         451899
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                  151
1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                 US-09-491-146A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15:
16:
17:
                                                                                                                                                                                                                                                                                                                     Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                              Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	Appl	Appli	Appl	pli	Appli	Appl	Appl	-	Appl	-		App	Appl	Appl	Appl
_	44,	1,	10,	l, Appl	8,	77,	14,	78,	46,	60,	152,	152	42,	44,	42,
Description	Sequence 44, App	Seguence 1, Appl	Sequence 10,	Sequence 1	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence
De				ഗ											
	8-44	7-1	8-10	-	2-8	1-17	8-14	US-09-921-397-78	8-46	8-60	6-152	1 - 152	6-42	6-44	1-42
	1-13	1-67	1-13	-308	6-87	1-39	1 - 13	1-39	1-13	1-13	9-04	8-28	9-04	9-04	8-28
	US-09-851-138-44	US-10-367-677-1	US-09-851-138-10	US-09-758-308-1	19-75	US-09-921-397-77	9-85	19-92	9-85	9-82	9-86	19-81	9-89	9-86	9-87
Ð	ns-(	US-	OS-	0S-0	-Sn	ns-(	ns-(	ns-(	ns-(	ns-(	ns-(	ns-(	ns-(	ns-(	OS-(
DB	10	15	70	σ	97	10	10	10	្ព	10	11	11	11	11	11
% Query Match Length DB	319	44	74	91	97	103	108	113	137	138	166	166	169	169	169
& Query Match	100.0	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4
Score	151	144	144	144	144	144	144	144	144	144	144	144	144	144	144
Result No.	-	7	æ	4	2	9	7	æ	6	10	11	12	13	14	15

1 US-09-878-281-44 Sequence 44, Appl 1 US-09-929-955-2 Sequence 2, Appl 1 US-09-1949-93 Sequence 2, Appl 1 US-09-1949-93 Sequence 3, Appl 1 US-09-194-913 Sequence 76, Appl 1 US-09-851-138-12 Sequence 12, Appl 1 US-09-811-138-12 Sequence 12, Appl 1 US-09-899-046-52 Sequence 52, Appl 1 US-09-899-046-14 Sequence 54, Appl 1 US-09-899-046-14 Sequence 54, Appl 1 US-09-878-281-54 Sequence 54, Appl 1 US-09-978-281-54 Sequence 54, Appl 1 US-09-978-281-54 Sequence 50, Appl 1 US-09-973-55 Sequence 50, Appl 1 US-09-973-55 Sequence 50, Appl 1 US-09-995-808-56 Sequence 50, Appl 1 US-09-995-808-56 Sequence 50, Appl 1 US-09-995-808-56 Sequence 23, Appl 1 US-09-95-55 Sequence 23, Appl 1 US-09-95-51 Sequence 24, Appl 1 US-09-95-51 Sequence 24, Appl 1 US-09-929-955 Sequence 24, Appl 1 US-09-929-955 Sequence 25, Appl 25-05 Sequence 25, Appl 25-05-2572-9 Sequence 27, Appl 26-05-05-05-05-05-05-05-05-05-05-05-05-05-	HEPATITIS C VIRUS GENOTYPES PROFHYLACTIC, THERAPEUTIC AND DIAGN SCII text output 38 .9 .7
978-9848-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-98888-988888-988888-9888888	ALIGNMENTS  38  UENCES OF H  IR USE AS P  SS/MS-DOS  SO/MS-DOS  Fd 6.0 / AS  709/851,138  11A A.  95775  91,775  R: INNS:004
US - 09 - 09 - 09 - 09 - 09 - 09 - 09 - 0	ALIGNME B51138 1 1 ERT VEN SEQUENCES THEIR USE NTS 7 7 WHITE & DU 433 FC-DOS/MS-D E Word 6.0 A: US/09/851 Y-2001 A: US/09/851 Y-2001 A: US/09/851 FD 948701 E P P 948701 E P P P P P P P P P P P P P P
	n US/09851131 183508A1 ENS, GEERT ENS, GEERT ENS, LIEVEN ON: NEW SEQUI AND THEID AGENTS CES: 207 AGENTS CES: 207 AGENTS CES: 207 AGENTS CES: 207 AGENTS COMPANION NON NON NON NON NON NON NON NON NON
1000 1000 1000 1000 1000 1000 1000 100	ALIC  44. APPLICATION US/09851138  42. INFORMATION: APPLICANT: MAERTENS, GEERT STUVER, LIEVEN TILLE OF INVENTION: NEW SEQUEN ADDRESSES: ADDRESSES: ADDRESSE: 207 CORRESPONDENCES: 207 CORRESPONDENCES: 207 TY: HOUSTON STATE: TEXAS COUNTRY: USA ZIP: 77210-4433 CIP: 77210-4433 COMPUTER READABLE FORM: MEDIUM TYPE: PLOPPY disk COMPUTER READABLE FORM: APPLICATION NUMBER: US/09, FILING DATE: 09-MAY-2001 PRIOR APPLICATION NUMBER: EP 946 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 946 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 946 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 946 FILING DATE: 28 JUN 1995 APPLICATION NUMBER: 29,77 REFERENCE/DOCKET NUMBER: 12,77 REFERENCE/DOCKET NUMBER: 29,77 REFERENCE/DOCKET NUMBER: 29,77 REFERENCE/DOCKET NUMBER: 29,77 REFERENCE/DOCKET NUMBER: 29,77
00000000000000000000000000000000000000	44 Application No. US20020. FORMATION: CANT: MAEKI CANT: MAEKI OF INVENTIG SPONDENCE A SPONDENCE A ADDRESSEE: P.O CITY: HOUST COUNTRY: USA COUNTRY COUNTR
	ALIGNM  109-811-138-44  109-811-138-44  GENERAL INFORMATION US/09851138  GENERAL INFORMATION: GERT STUTVYER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES AND THEIR USE  NUMBER OF SEQUENCES: 207  CORRESPONDENCE ADDRESS: CORRESPONDENCE ANDRESS: STREET: P.O. BOX 4433  STREET: P.O. BOX 4433  STREET: P.O. BOX 4433  COUNTRY: HOUSTON STATE: TEXAS COUNTRY: USA INTORNATION TYPE: FLOPPY disk COMPUTER: IBM PC COMPATA: APPLICATION DATA: APPLICATION DATA: APPLICATION DATA: APPLICATION NUMBER: US/09/85 FILING DATE: 31 Oct 1994 APPLICATION NUMBER: EP 94870 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 95870 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 95870 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 95870 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 95870 FILING DATE: 21 Oct 1994 APPLICATION NUMBER: EP 95870 FILING DATE: 1 NUMBER: INN NAME: KAMMERER: NUMBER: INN NAME: KAMMERER: NUMBER: INN NAME: KAMMERER: NUMBER: INN NAME: KAMMERER: NUMBER: INN NAME: REFERENCE/DOCKET NUMBER: INN
UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU	RESULT GENT GENT GENT GENT GENT GENT GENT GEN
•	

```
Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION.
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: PROTEINS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.034902
CURRENT APPLICATION NUMBER: US/09/758,308
PRIOR FILING DATE: 2001-01-10
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.0
ENGTH NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                               CURRENTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 6.0 / ASCII text output
CURRENT APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 OCt 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 21 OCt 1994
APPLICATION NUMBER: EP 95870076.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95.4%; Score 144; DB 10;
96.4%; Pred. No. 2.2e-13;
vative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 144; DB 9;
Pred. No. 2.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: KAMMERER, PATRICIA A.
RESISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ 1D NO: 10:
                                            WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                    ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                            STREET: P.O. BOX 4433
CITY: HOUSTON
  NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                            ADDRESSEE: ARNOLD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95.4%;
ilarity 96.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                                                                                                    STATE: TEXAS
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: MAERTENS, GEERT STUYVER, LIEVEN STUYVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 2003-02-19
PRIOR FILING DATE: 1099-09-07.
PRIOR FILING DATE: 1099-09-07.
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
TENGRU: AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                             Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 44;
                                                                                                                                                                                                     Indels
                                                                                                                                                             DB 10;
                                                                                                                                                         100.0%; Score 151; DB 10;
dlarity 100.0%; Pred. No. 1.1e-13;
Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 144; DB 15;
Pred. No. 1.3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                            ; TOPOLOGY: linear; MOLECULE TYPE: peptide; SEQUENCE DESCRIPTION: SEQ ID NO: 44: US-09-851-138-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                  LENGTH: 319 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/10367677 ublication No. US20030118604A1 ENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: JOLIVET, MICHEL
APPLICANT: PENIN, FRANCOIS
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
Best Local Similarity 90...
Best Local Similarity 90...
27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAGES: 3392-3396
DATE: 1991
; RELEVANT RESIDUES: 2 TO 45
US-10-367-677-1
                                                                                                                                                         Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                          RESULT 2
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

```
APPLICANT: MAERTENS, GEERT STUYVER, LIEVEN STUYVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRUTES: IBM PC compatible
CORREATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICTOSOFT WORD 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 40HANDAM:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 108;
                                                     Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 144; DB 10;
Pred. No. 3.3e-13;
                                                  Score 144; DB 10;
Pred. No. 3.2e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                        Sequence 14, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Sequence 78, Application US/09921397
; Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                         AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 95.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative (
                                                    Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                         US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-851-138-14
                                                                                                                                                                                                                    RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 97;
                                                         DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95.4%; Score 144; DB 10
96.4%; Pred. No. 3e-13;
tive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 77, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVERTION: SID nucleic acids and pol:
TITLE OF INVERTION: pathogenic strain of the
TITLE OF INVERTION: applications thereof
FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE: 28-F8B-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808-157A
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
SECUENCE CHARACTERISTICS:
SECUENCE CHARACTERISTICS:
SECUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 95.4
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 77
LENGTH: 103
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-921-397-77
                                                     35-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
```

```
APPLICANT: MAERTENS, GEERT
APPLICANT: MAERTENS, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Microsoft Word 6.0 / ASCII text output CURRENT APPLICATION DATA:
                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 144; DB 10;
Pred. No. 4.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                 Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 138 amino acids
                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95.4%;
                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 96.4
Matches 27; Conservative
               27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
US-09-899-046-152
                                                                                                                                                                                US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-60
               Matches
                                                                                           g
                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AFFLICANT. STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                        ö
             polypeptides selected from a the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: B 94870166.9
FILING DATE: CIOKNOWN>
APPLICATION NUMBER: B 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 22 98 Jun 1995
                                                                                                                                                                                                                                                                                                                                               Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 137
                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                          Score 144; DB 10;
Pred. No. 3.5e-13;
); Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 144; DB 10;
Pred. No. 4.3e-13;
TITLE OF INVENTION: SID nucleic acids and por TITLE OF INVENTION: pathogenic strain of th TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ CURRENT FILING DATE: 2001-08-02 FRIOR APPLICATION NUMBER: EP 00402225.7 PRIOR FILING DATE: 2000-08-03 NUMBER OF SED ID NOS: 156 SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: INNS:004 INFORMATION FOR SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKTKRNTIRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SID nucleic acids and pathogenic strain of applications thereof
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 46, Application US/09851138 Publication No. US20020183508A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD,
                                                                                                                                                                                                                                                          ; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 9
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-851-138-46
                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3
```

```
Floppy disk
                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;;
0
                                                                                                                                                                                                                                                                                                                                      Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 144; DB 11; Length 166;
Pred. No. 5.3e-13;
0; Mismatches 1; Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PE-ENTING PROS #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
APPLICATION DATA:
APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylax NUMBER OF SEQUENCE: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                      DB 11;
                                                                                                                                                                                                                                                                                                                                    Query Match 95.4%; Score 144; DB 11; Best Local Similarity 96.4%; Pred. No. 5.3e-13; Matches 27; Conservative 0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                  FRPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Sequence 42, Application US/09899046
; Publication No. US20030008274A1
                                                                                                                                                    FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27; Conservative
                                                                                                                                                                                                                                                   TOPOLOGY: linear;
MOLECULE TYPE: protein
US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION:
TITLE OF INVENTION: 94
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 27; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                      Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                               Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 144; DB 11;
Pred. No. 5.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                  Score 144; DB 11;
Pred. No. 5.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKKNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/09/899,046
                                                                                                                        PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Sequence 44, Application US/09899046; Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 42, Application US/09878281; Publication No. US20030032005A1 GENERAL INFORMATION:
APPLICANT: TITLE OF INVENTION: New sequence: TITLE OF INVENTION: qenotypes for NUMBER OF SEQUENCES: 270
                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 42: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION FOR SEQ ID NO: 44: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                    95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95.4%;
                                                                                                                                                                                                                          LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      169 amino acids
                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 96.49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 95.48
Best Local Similarity 96.48
Matches 27; Conservative
                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
```

Search completed: August 7, 2003, 12:01:12 Job time : 15.3636 secs

us-09-491-146a-24.rapb

```
AFFLICANT. STUTYER, LIEVEN
STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                               Gaps
                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 108;
                                                       Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 149; DB 10;
Pred. No. 1.9e-13;
1; Mismatches 0;
                                                       Score 149; DB 10;
Pred. No. 1.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: <unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                    18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET: P.O. BOX 4433
CITY: HOUSTON
STRIE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                               Sequence 14, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                           APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                    AGENTS
                                                       97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 97.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD,
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                 Query Match
Best Local Similarity 96.4<sup>1</sup>
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
US-09-921-397-78
                                                                                                                                                                                                                                                JS-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-851-138-14
                                                                                                                                  ò
                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 97;
                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 149; DB 10;
Pred. No. 1.7e-13;
1; Mismatches 0;
                                                                                                     GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TILLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: SALITE 701-E Columbia Square
STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR PILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENGTH: 103
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: HYBRIGENICS
TITLE OF INVENTION: 510 nucleic acids and printe OF INVENTION: 510 nucleic acids and printe OF INVENTION: pathogenic strain of TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILLING DATE: 2001-08-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPOROTKRNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                               ZIP: 2004
COMPUTER REARABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 77, Application US/09921397 Patent No. US20020151484A1
                                                                 Sequence 8, Application US/09756875
Patent No. US20020150990A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                                 D. C.
                                                                                                                                                                                                                                                           STATE: D
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
OF HEPATITIS C VIRUS GENOTYPES
AS PROPEYLACTIC, THERAPEUTIC AND DIAGNO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
             ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 149; DB 10;
Pred. No. 2.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
FILING DATE: <1004 1994
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                    1 PKPOROTKRNTNRRPQDVKFPGGGQIVG 28
                                                                             1 PKPORQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: NEW SEQUENCES
AND THEIR USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: FIDPPY disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 152, Application US/09899046 Publication No. US20030008274A1
                                                                                                                                                                                                  Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 138 amino acids
             ä
                                                                                                                                                                                                                                                                  APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                      AGENTS
                                                                                                                                                                                                                                                                                        STUYVER, LIEVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 96.47
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: amino acid
           27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
US-09-899-046-152
                                                                                                                                                        RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-851-138-60
             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                    ò
                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: MAERTENS, GEERT STUVER, LIEVEN STUYVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                           Gaps
        polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
CITY: HOUSTON
STATE: TEXAS
COMPUTR: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION NDATA:
APPLICATION NDATA:
APPLICATION NUMBER: US/99/851,138
FILING DATE: 09-MBY-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                 Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 137;
                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                               DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 149; DB 10;
Pred. No. 2.5e-13;
                                                                                                                                                                                                                                                                                                                                             Score 149; DB 10
Pred. No. 2e-13;
1; Mismatches
TITLE OF INVENTION: SID nucleic acids and polification of the TITLE OF INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ CURRENT APPLICATION NUMBER: US/09/921,397 CURRENT FILING DATE: 2001-08-02 FRIOR APPLICATION NUMBER: EP 00402225.7 PRIOR FILING DATE: 2000-08-03 NUMBER OF SEQ ID NOS: 156 SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: «Unknown»
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: INNS:004 INFORMATION FOR SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: 11near

MOLECULE TYPE: peptide .

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-851-138-46; Sequence 46, Application US/09851138; Publication No. US/0020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 137 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97.48;
                                                                                                                                                                                                                                                                                                                                             ch 97.4%;
1 Similarity 96.4%;
27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                        ; ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                          LENGTH:
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
NUMBER OF SEQUENCES: 270
NUMBER PREDABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PAtentIN Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 169;
                                                                                                                                                                                                                                                                                                                                                             DB 11; Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                    COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; DB 11; 1
3.1e-13;
                                                                                                                                                                                                                                                                                                                                                         Score 149; DB 11;
Pred. No. 3.1e-13;
1; Mismatches C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 149; DB; Pred. No. 3.1e-1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKPORKTKRNTNRRPQDVKFPGGGOIVG 32
                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 44, Application US/09899046; Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Sequence 42, Application US/09878281; Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97.4%;
                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
        Floppy disk
                                                                                                                                                                                                                                        LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity 96.49
Matches 27; Conservative
                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                              ) MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 166;
                                                                                                                                                                                                                                                                                                                                       Length 166;
                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIDLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PARIATIN RELEASE #1.0, Version #1.25 (EPO)
CURREYM APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                       DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 11;
                                                                                                                                                                                                                                                                                                                                     Score 149; DB 11
Pred. No. 3e-13;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 149; DB 11
Pred. No. 3e-13;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 152, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 42, Application US/09899046
Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                   PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                       97.4%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 97.4%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                     i: 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                         FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                        TOPOLOGY: Ilnear
MOLECULE TYPE: protein
US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-899-046-42
                                                                                                                                                                                                                 LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                   >,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                    음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
us-09-491-146a-24.rapb
```

```
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBA PC compatible
COMPUTER: IBA PC compatible
COMPUTER: IBA PC compatible
CORRENT APPLICATION DATA:
FILING DATE: VS/09/878,281
FRICATION NUMBER: US/09/878,281
FRICATION NUMBER: US/09/878,281
FRICATION FOR SEO ID NO: 42:
FILING DATE: APPLICATION POR SEO ID NO: 42:
FILING DATE: APPLICATION FOR SEO ID NO: 42:
FILING
```

Search completed: August 7, 2003, 12:01:12 Job time: 14.3636 secs

core-envelope core-envelope

Scoring table:

Searched:

Minimum DB Maximum DB

Database

Perfect score:

Sequence:

OM protein

Run on:

```
Non-A, Non-B Hepat
Peptide VIIIE base
Anti-HCV antibody
Prototype peptide
HCV core protein p
PT-NANB VIE2 Stru
Hepatitis C virus
Antigen pHCb101.
Antigen pHCb101.
HCV fragment 1 / I
HCV fragment 2 / I
Branched epptide H
Branched epptide H
HCV capsid peptide
Non-A, non-B hepat
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Human hepatitis C
Human hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
HCV bait polypepti
HCV capsid core pr
HCV core-envelope
                                                                                                                                                                                                                                                                                                              Antigenic epitope
HCV core protein N
                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
HCV bait polypepti
Protein encoded by
                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
Protein encoded by
                                                                                    HCV core-envelope
HCV core-envelope
                                                                                                                                      HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NC mosaic protein amino acid fragment K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                    AAW66083
AAY26952
AAY94410
AAY94409
AAR29534
                                                                                                                                                                                                                                                                                                AAR66619
                                                                                                                                                                                                                                                                                                              AAB31695
                                                                                                                                                                                                                                                                                                                                                                                                                         AAR98348
                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR29530
                                                                                                                                                                                                                                                                                       AAR51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY06683 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fields HA, Khudyakov YE;
 WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
WO9910506-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-MAR-1999
AAY06683;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 1
 NC mosaic protein
NC mosaic protein
Heparitis C virus
NC mosaic protein
Blood transmiscibl
HCV-S1 full-length
CN14 fragment of H
HCV capsid peptide
                                                                                Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                  | SIDSI/gcddata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcddata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcddata/geneseq-geneseqp-embl/AA1981.DAT:*
| SIDSI/gcddata/geneseq-geneseqp-embl/AA1981.DAT:*
| SIDSI/gcddata/geneseqp-embl/AA1982.DAT:*
| SIDSI/gcddata/geneseqp-embl/AA1992.DAT:*
| SIDSI/gcddata/geneseqp-embl/AA1992.DAT:*
| SIDSI/gcddata/geneseqg-embl/AA1992.DAT:*
| SIDSI/gcddata/geneseqg-embl/AA2002.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
         GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                          tal number of hits satisfying chosen parameters:
                                                                                                                                            153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                     1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                        protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY06683
AAY06678
AAR58593
AAY06675
AAR53417
AAR30687
AAR30687
                                                                                                                                                                                              Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                        A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                     seq length: 0
seq length: 200000000
                                                                                                                                 US-09-491-146A-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    146335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3010
3010
3010
36
38
                                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0
```

Score

Result Š

```
artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAV06673-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                              FILILITE FILITIFILITIFILITE PREPARENCE PREPARENCE PROPERTY OF THE PROPERTY OF 
                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR58593 standard; Protein; 312 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 1; Page 18; 22pp; Japanese
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY06675 standard; Protein; 28
                                                                                                                                                                                                                                                               Query Match 96.7%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92JP-0207391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    93JP-0193104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (KOKU-) KOKUSAI SHIYAKU KK.
(SANW ) SANWA KAGAKU KENKYI
(TOFU ) TONEN CORP.
(TOKR-) ZH TOKYOTO RINSHO I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09-MAY-1995 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1994-298800/37.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               312 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                     28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAQ70543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JP06225770-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUL-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-AUG-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR58593;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 4
AAY06675
ID AAY0
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
   8×3333338
                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME OF COLOR OF THE PROPERTY 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                     The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis of the method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein mosaic gene and protein is also useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 153; DB 20;
Pred. No. 2.6e-14;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          peptides from different genotypes of a species detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKTKRKAHRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NC mosaic protein amino acid fragment F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY06678 standard; Protein; 28 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   comprising the NC mosaic protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 5; Fig 9; 66pp; English.
                           Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   uery Match similarity 100.vv-st Local Similarity 100.vv-rac 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY06678;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
AAY06678
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAQ70543 is a fragment of hepatits C virus (HCV) or non-A non-B hepatits virus (NANBH) core and envelope one structural regions, encompassing base pairs 349-1285. It codes for AARS8593 an antigen to a structural region of the HCV virus, which can be used in the diagnosis of NANBH patients and the detection of HCV carriers.
                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 312;
  Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for diagnosis of NANBH and detection of HCV
                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
Score 148; DB 20;
Pred. No. 1.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 141; DB 15;
Pred. No. 1.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TONEN CORP. ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus 349-1285 fragment antigen.
                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                       28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 PKPQRKTKRNANRRPQDVKFPGGGQIVG 29
```

Gly

Leu

17-JUN-1999

AAY06675;

WO9910506-A1

04-MAR-1999

21-AUG-1998; 25-AUG-1997;

Claim

```
Location/Qualifiers
                                                                                                                                                                                                          label- Ser, Arg,
                                                                                                                                                                                                                                                                     'label Gln, Pro,
                                                                                                                                                                             /label= Gly, Asp
                             label- His, Arg
                                                                                                                     label- Asp, Asn
                                                                                                                                                                                                                                      label- Thr, Ala
                                                                                                                                                                                                                                                                                                 'label- Lys, Arg
                                                                                                                                                                                                                                                                                                                                                          'label- Asn, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label- Ser, Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           abel- Thr, Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Glu, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           abel- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Met, Thr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             abel= Asn, Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Asn, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label- Gly, Ile
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label= Cys, Arg
                                                          /label= cys,
                                                                                        label- Leu,
                                                                                                                                                                                                                                                                                                                            label- Gly,
                                                                                                                                                  'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                       label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                    .abel- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        abel- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 2488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 1455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 2844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2968
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 1017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 120]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 2862
                                                                                                                                                                                         Misc-difference 398
                                                                                                                                                                                                                                                                                Misc-difference 410
                                                                                                                                                                                                                                                                                                                                         Misc-difference 430
                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 438
                                                                        Misc-difference 246
                                                                                                    Misc-difference 263
                                                                                                                                 Misc-difference 291
                                                                                                                                                                                                                      Misc-difference 400
                                                                                                                                                             Misc-difference 31
                                                                                                                                                                                                                                                    Misc-difference 40
                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 47
   Key
Misc-difference 22
                                                                                                                                                                                                                                                                                                             Misc-difference 41
                                           Misc-difference 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 29
                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also provided. The method is designated restriction endonuclease assisted rigation (REAL). The mosalc protein and the artificial mosalc protein are useful for detecting a hepatitis infection in an individual. The mosalc gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the infection, especially hepatitis C. The method of synthesizing the spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NS5 region.
                                                                              Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                        The invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                              New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90.2%; Score 138; DB 20; Length 2 89.3%; Pred. No. 3.2e-12; Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NC mosaic protein amino acid fragment C.
                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                           5; Fig 9; 66pp; English
                                                                                                                                                                                                                  98WO-US17385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Non-A, non-B hepatitis virus
                                                                                                                                                                                                                                            97US-0921887
                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-JAN-1995 (first entry)
                                                                                                                                                                                                                                                                                                       Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Jest Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                   WPI; 1999-204671/17.
                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28 AA;
```

Sequence

wery Match

ò a AAR53417;

RESULT 5 AAR53417

```
The invention relates to nucleic acid construct which comprises an expression cassette including a first polynuclectide region including a 5 non-cooling region (URN) sequence of an RNA, virus and at least an N-terminal portion of a coding sequence of RNA virus, a second colynuclectide region including a 3' untranslated region (UTN) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a third polynuclectide region encoding a reporter molecule, flanked by first and second polynuclectide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug resistance of an RNA virus. The present sequence is Hepatlitis C virus (HCV) isolate HCV-Si full-length polyprotein.
  Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of HCV infection and to raise antibodies against it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Detection of hepatitis C virus - using oligopeptide fragment of core region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90.2%; Score 138; DB 23; 89.3%; Pred, No. 4.1e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                             Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAW06487 standard; Protein; 36 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CN14 fragment of HCV core region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93JP-0156026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (IMMO ) IMMUNO JAPAN KK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1995-047903/07
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAT45055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAW06487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           antibody.
                                                                                   cassette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW06487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were Cl00 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                    non-A non-B hepatitis virus DNA - used for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 15; Length 3010;
Pred. No. 4.1e-10;
1; Mismatches 2; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MOLE-) INST MOLECULAR & CELL BIOLOGY (EHRL/) EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 1; Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       М.
                            /label- Tyr, Cys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hong
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90.2%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                            detection of hepatitis virus
                                                                                                                                                                            92JP-0051885
                                                                                                                                                                                                                             92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-JUL-2000; 2000US-220248P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 38, 89 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUL-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV-S1; Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fan YH, Lim SP, Lim
Misc-difference 2990
                                                                                                                                                                                                                                                                                                                                                                                                    Blood-transmissible
                                                                                                                                                                                                                                                                                                                           WPI; 1994-163130/20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-280605/32.
N-PSDB; AAD33038.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3010 AA;
                                                                                                                                                                                                                                                                            (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAQ63499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200208447-A2.
                                                                         JP06105690-A
                                                                                                                                                                            10-MAR-1992;
                                                                                                                                                                                                                           10-MAR-1992;
                                                                                                                          19-APR-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31-JAN-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAE20477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAE20477
ID AAE;
XX
AC AAE;
XX
DT 01-,
388888888888
```

g

ò

ö

Gaps

Fri Aug

셤

ò

```
Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                            Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR30689 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                 (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative 1
                                                                                                                                                                                                                                                                        91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                92WO-US03635
                                      (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV capsid peptide No. 25.
                                                                              HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                                                                            Jolley ME, Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                      WPI; 1993-018073/02
                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38 AA;
                                                                                                                                                                                        W09222571-A1.
                                                                                                                                                                                                                                              29-APR-1992;
                                                                                                                                                                                                                                                                         13-JUN-1991;
                                                                                                                                                                                                                                                                                      20-JUN-1991;
                                      25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                    23-DEC-1992.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          antibodies
            AAR30688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR30689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR30689
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The sequences given in AAR30665-89 represent fragments of the hepetitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope artibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                   Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
           Length 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 38
                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 137; DB 14;
Pred. No. 6.1e-12;
1; Mismatches 2;
          Score 137; DB 16;
Pred. No. 5.8e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Updated on 25-MAR-2003 to correct PN field.)
                                                                              1 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR30688 standard; peptide; 38 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                           AAR30687 standard; peptide; 38 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BAXT ) BAXTER DIAGNOSTICS INC.
          89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                     91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                            92WO-US03635
                                                                                                                                                                                                                  (updated)
(first entry)
                                                                                                                                                                                                                                                          HCV capsid peptide No. 23
                                     25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1993-018073/02
           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          vīrus – for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                   W09222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                          29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-JUN-1991;
                                                                                                                                                                                                                25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jolley ME,
                                                                                                                                                                                      AAR30687;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seguence
                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR30688
ID AAR3
                                                                                                                                             AAR30687
ID AAR3
XX
AC AAR3
XX
XX
DT 25-1
DT 11-1
XX
HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                  RESULT 8
```

```
Gaps
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 3M and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope artibodies in the diagnosis of non-A, non-B hepatitis (NANDH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 137; DB 14; Length 38;
Pred, No. 6.1e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                        (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
```

a

ð

Gaps

}8888888888**%** 

ద à

```
The present sequence represents a Hepatitis C virus (HCV) protein sequence from the disclosure of the present specification. The present specification describes a chimeric HCV peptide antigen which comprises at least 2 peptide epitope regions from the HCV polypeptide core region, 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS4 region. The antigen binds specifically with an antibody produced by a human infected by HCV. The peptide can detect a wide range of HCV infections with high sensitivity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New chimeric peptide antigen derived from hepatitis C virus protein - useful for detecting \mbox{HCV} infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus; HCV; chimeric; antigen; detection; core region; epitope; NS3; NS4; infection.
                                                                                                                                                                                                                                         ö
                                                                                              Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69).
Both genes contain the core, ENV, NS1, NS2 and NS3 regions.
A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                 non-B hepatitis virus antigen - useful for
                                                                                                                                                                                                        Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                  Score 137; DB 15;
Pred. No. 6.1e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                       1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Page 11; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Page 24; 30pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus C-1 protein 1-43.
                                                                                                                                                                                                                                                                                                                                                                                     AAW37380 standard; Protein; 43
                                                                                                                                                                                                 89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97JP-0027015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96JP-0024045
              DNA coding a Non-A, non-B }
detecting HCV within serum
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                             Query Match
Best Local Similarity 89.3 Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 89.3 es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1998-022248/03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TOFU ) TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis c virus.
                                                                                                                                                                 38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JP09278794-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        11-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09-FEB-1996;
                                                                                                                                                                   Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                     AAW37380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                       AAW37380
ò
                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                         The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contained in the capsid protein of the virus and themselves contained in acroups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                         Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; core; ENV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Non-A, non-B hepatitis virus corel region fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 137; DB 14;
Pred. No. 6.1e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                           Disclosure, Fig 1F; 66pp; English,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SANW ) SANWA KAGAKU KENKYUSHO CO
                                                                                                                                                        (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AARS4065 standard; Protein; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.5%;
                                                                         92WO-US03635.
                                                                                                     91US-0714471.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                          Leahy DC,
                                                                                                                                                                                                                           WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1994-205028/25.
N-PSDB; AA064067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TONEN CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38 AA;
                                                                                                     13-JUN-1991;
20-JUN-1991;
       W09222571-A1
                                                                      29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP06141870-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12-MAR-1992;
                                      23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-FEB-1995
                                                                                                                                                                                          Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-MAY-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               antibodies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR54065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AARS4065
ID AARS
XX
```

ö

Gaps

```
Peptides AAY26949-Y26955 represent peptide epitopes derived from the N-terminus of the hepatitis C virus core protein. The peptides are used to generate monoclonal antibodies or antibody fragments specific for hepatitis C virus (HCV) core protein. The monoclonal antibodies are used for early diagnosis of HCV infections, especially by two-antibody
                                                                                                                                                                                                                                                                                                                                                           Monoclonal antibodies useful for detecting and/or quantifying hepatitis \mathsf{C} virus core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polypeptides that bind to anti-hepatitis C virus antibodies, useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human hepatitis C virus core protein N-terminus, residues 2-45.
                     Epitope; hepatitis C virus; core protein; moncclonal antibody;
diagnosis; infection; sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΰ
                                                                                                                                                                                                                                                                                           Jolivet M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Jolivet-Reynaud
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 137; DB 20;
Pred. No. 7.1e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                         Paranhos BC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Æ
                                                                                                                                                                                                                                                                                                                                                                                                             Claim 3; Page 10; 19pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY94410 standard; peptide; 44
                                                                                                                                                                                                                                                                                         Yvon S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 89.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                  98FR-0003087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99WO-IB01933.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0196155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ۵,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dalbon
                                                                                                                                                                                                                                                                                       Piga N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sandwich immunoassay
                                                                                                                                                                                                                                                                                                                        WPI; 1999-530397/45
                                                                                                                                                                                                                                                     (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-411934/35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (INMR ) BIO MERIEUX
                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200031130-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 н
                                                                                                           FR2775690-A1
                                                                                                                                                                                                                  09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-NOV-1998;
                                                                                                                                            10-SEP-1999
                                                                                                                                                                                                                                                                                       Jolivet RC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      02-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY94410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY94410
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   complex formation). Antibodies are used correspondingly to detect HCV or related antigens. The peptide and antibodies may also be used to treat or prevent HCV infections. The present sequence represents the N-terminus of the core p21 protein of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to a peptide, which is recognised by antibodies against amino acids 2-45 at the N-terminus of the core (or nucleocapsid) by 1 protein of hepatitis C virus (HCV), or its variants. The peptide has a tertiary structure consisting of two alpha-helical fragments, almost perpendicular to each other in space, connected by a junction peptide. Excluded are all proteins and peptides comprising, or consisting of, the N-terminal part of p21 (starting from amino acid 1 or 2). Also new are (1) monoclonal or polyclonal antibodies produced using the peptide as an immunogen and (2) complex consisting of the peptide specifically bound to some other molecule, particularly peptide or nucleotide fragments or functionalised aromatic compounds. The peptide is used (1) as immunogen for generating antibodies and (11) for detecting and guantifying either antibodies against p21 or HCV-derived mRNA (by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       nsed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New peptide from the N-terminus of hepatitis C virus p21 protein containing the immunodominant epitope - and related antibodies, ufor diagnosis, treatment and prevention of hepatitis C infection
                                                                                                                                                                          antigenic; nucleocapsid; p21 protein; alpha-helical; immunogen; antibody; hepatitis C virus; epitope; N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                               Penin F;
                                                                                                                                         Hepatitis C virus p21 protein N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.5%; Score 137; DB 19;
89.3%; Pred. No. 7.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus Core protein amino acids 2-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Ladaviere L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. 7.16
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW66083 standard; peptide; 44 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Page 16; 37pp; French
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                              Lacoux X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY26952 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                     98WO-FR00442
                                                                                                                                                                                                                                                                                                                                                                      97FR-0002878
                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-DEC-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                            Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                         (INMR ) BIO MERIEUX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1998-495793/42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
tes 25; Conserv
                                                                                                                                                                                                                             hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44 AA;
                                                                                                    16-NOV-1998
                                                                                                                                                                                                                                                              WO9839360-A1
                                                                                                                                                                                                                                                                                                                                   05-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                      05-MAR-1997;
                                                                                                                                                                                                                                                                                                  11-SEP-1998.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Dalbon P,
                                                                   AAW66083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY26952;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
RESULT 1
AAW66083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY26952
ID AAY2
XX
AC AAY2
XX
DT 21-D
XX
DE Hepa
```

SXSS

ð g

```
diagnosing and preventing hepatitis C infections -
```

Claim 1; Page 42; 50pp; English.

The present peptide, designated 542G, corresponds to residues 2 to 45 of the N-terminus of the human hepatitis C virus (HVV) core protein. It is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoreactivity with all the sera of individuals or blood samples infected with HCV and which possess substituted for homologous amino acros protein. An amino acid may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino acids, amine groups may be acetytated, and so on. The native antigenic sequence and its antigenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against the virus. 

44 AA; Sequence

ô Gaps ö 89.5%; Score 137; DB 21; Length 44; 89.3%; Pred. No. 7.1e-12; Live 1; Mismatches 2; Indels Query Match
Best Local Similarity 89.34
Atches 25; Conservative

1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28

ĩ q

4 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 31

Search completed: August 7, 2003, 11:14:10 Job time : 38.6364 secs

```
89.5
137
30
                                                                                                                         August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
               GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                          tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                     US-09-491-146A-33
153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                        283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                          OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length; 0
Maximum DB seq length; 2000000000
                                                                                                                                                                                                     Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                        Searched:
                                                                                                                               Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Database :

## SUMMARIES

	Description	genome polyprotein																		genome polyprotein		pro	polyprotein - hepa	tein -	genome polyprotein	genome polyprotein	genome polyprotein	genome polyprotein	hypothetical prote
		1 1 1 1 1 1																											
KIES		! ! ! !																											
SUMMAKIES	ID	A45573	S41353	S41355	S41357	S41348	S41371	S41341	S41370	S41369	S41368	S41342	541344	S41350	541354	S41345	S41347	S41343	841346	S21471	S12707	PC1284	JQ1925	JQ1926	JH0711	S19876	S18031	S18032	PN0677
	DB		N	~	~	7	~	~	~	N	~	7	7	~	~	7	~	~	~	ď	7	~	7	7	~	7	7	N	~
	Length	3010	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115	115	118	369	441	513	520	523	550	782	782	782	787
æ	Query Match	90.2	89.5	89.5	89.5	89.5	89.5	89.5		89.5			89.5		89.5		89.5	89.5	89.5		•		89.5	89.5				89.5	89
	Score	138	137	137	137	137	137	137	137	137	137	137	137	137	137	137	137	137	13.7	137	137	137	137	137	137	137	137	137	137
	Result No.	٦	~	m	4	S	9	7	8	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	58

9enome polyprotein PC2219 PC2219 GNWYTC GNWYCJ GNWYCJ GNWYCJ GNWYCJ GNWYCH GENOME POLYProtein GNWYCH GNWYCH GENOME POLYProtein GNWYCH GENOME POLYProtein GENOME POLYProtein GENOME POLYProtein GENOME POLYProtein GENOME POLYProtein GENWYCH GENOME POLYProtein GENWYCH S41356 GENOME POLYProtein	RESULT 1  AASTOLT 1  A	<pre>sion 26-Jul-1996 #text_change 17-Nov-2000 Brouwer, J.T.</pre>
137 89.5 874 137 89.5 876 137 89.5 3010 137 89.5 3010 137 89.5 3011 137 89.5 3011 137 89.5 3011 137 89.5 3011 137 89.5 108 133 86.9 108 133 86.9 114 133 86.9 114	RESULT 1  A45573 genome polyprotein - hepatitis C virus (st. N; Contains: capsid protein C; envelope protein NS4; nonstructural protein NS4b; protein NS4b; cylus C; Date: 19-May-2000 #sequence_revision 19-15. Species: hepatitis C virus C; Date: 19-May-2000 #sequence_revision 19-15. Richard, T; Kato, N; Nakagawa, M; Oots N; Tanka, T; Kato, N; Nakagawa, M; Oots N; Tanka, T; Rato, N; Nakagawa, M; Oots N; Tanka, T; Tanka, T; Rato, N; Nakagawa, M; Oots N; Tanka, T; Protein C did not type: DNA A; Reference number: A45573; MuID: 92295714; A; Retence number: A45573 MuID: 92295714; A; Retence number: A45573 MuID: 92295714; A; Residues: 1-910 C-TAN A; Residues: 1-90-1237/Region: nucleotide-binding mot F; 101-1317/Region: DEXH motif E-1010 C-TAN Matches 25; Conservative 1; Misma Qy Nebels: Result 2  RESULT 2  RESULT 2  S41333  S41333  S41333  S41333  S41333  S41333  S41333  SAPPER HERRINTRRPODVKFPGGGOIVG N; Conteains: core protein C; virus (ge N; Conteains: core protein C; virus (ge N; Repeties: hepatitis C virus (ge N; Represion: hepatitis C virus (ge N; Represion: hepatitis C virus (ge N; Represion	, # S
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	RESULT 1 A45573  94000me polyproj N; Contains: caj N; Contains: caj C; Date: 19-May. C; Jacession: A; R; Tanaka, T;; J; Virus Res. 23, A; Title: Molecule A; Reference nu A; Accession: A, A; Reference nu A; Accession: A, A; Reference nu A; Accession: A, A; Residues: 1:- A; Note: sequence A; Residues: 1:- A; Note: sequence C; Superfamily: C; Keywords: A; Product F; 192-389/Product F; 192-389/	A; variety C; Date: 1 C; Accessi R; van Doc

N

ö

Gaps

ö

Indels

#text\_change 17-Nov-2000

g ŏ

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 5, N5
A; Variety: genotype 5, N5
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41371
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by Liph.
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                          C) Accession: S41348
R) van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A) Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A) Reference number: S41341
                                                                                                                                                                                                                                                         polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 2; Length 112;
Pred. No. 1.6e-11;
1; Mismatches 2; Indels
  Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A.Experimental source: genotype 1, N6
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F.1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 137; DB 2; I
Pred. No. 1.6e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A:Experimental source: genotype 5, N5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                             N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 1, N6
C.Date: 19-May 1994 #sequence_revision 26-Jul-1996
                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPORKTKRKAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89.5%;
Similarity 89.3%;
5; Conservative 3
                   89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
A;Cross-references: EMBL:229451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Residues: 1-112 <VAN>
A;Cross-references: EMBL:Z29474
                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA A; Residues: 1-112 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatit.
N;Contains: core protein
C;Species: hepatitis C virus
                      Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S41371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Simi
Matches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
                                                                                                                                                                                                                                                              депоше
                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Species: hepatitis C virus (yemotype 2, N4) (inayment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41355
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Bescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41356
A;Rolecule type: genomic RNA
A;Residues: 1-108 <VAN>
A;Residues: 1-108 <VAN>
A;Residues: Gunonic RNA
A;Experimental source: genotype 2, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Date: 19-May 1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                     5 by LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LIPA
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ρ
                                                                                                                                                                                                                                                                                                      Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108;
                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     to 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
ដ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C. Accession: 34.1357
R. van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R. van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
B. beacription: Analysis of hepatitis C virus genotypes 1
A. Reference number: S41341
A. Accession: S41357
A. Molecule type: genomic RNA
A. Residues: 1-108 <-VANA
A. Residues: 1-108 <-VANA
A. Cross-references: EMBL: Z29460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 89.5%; Score 137; DB 2; L. Best Local Similarity 89.3%; Pred. No. 1.6e-11; Matches 25; Conservative 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                 Score 137; DB 2; I
Pred. No. 1.6e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A.Experimental source: genotype 2, N6
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F.1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB
                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKKKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 137;
                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
```

RESULT 4

ð

ö

Gaps

ö

Indels

Length 108;

ö

Gaps

m

```
A; Accession: S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S41342
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
A;Variety: genotype 1, N1
C;Date: 19 May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41341
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the Embl. Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Roccession: S41341
A;Molecule type: genomic RNA
A;Rolecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N3
C;Variety: genotype 5, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C:Accession: S41370
R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A:Reference number: S41341
A:Accession: S41370
A:Residues: 1-114 cVAN>
A:Cross_references: EMBL:229473; NID:9443908; PIDN:CAA82611.1; PID:9443909
A:Experimental source: genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: genomic RNA
A; Residues: 1-114 «VAN>
A; Cross-references: EMBL:229472; NID:g443906; PIDN:CAA82610.1; PID:g443907
A; Experimental source: genotype 5, N3
C; Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                    A; Cross - references: EMBL: Z29444; NID: 9443850; PIDN: CAA82582.1; PID: 9443851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S4134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 137; DB 2; Length 112;
Pred. No. 1.6e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 137; DB 2; Length 114;
Pred. No. 1.6e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                              A; Experimental source: genotypel, N1 (S) Superfamily: hepatitis C virus genome polyprotein C; Reywords: capaid protein; core protein; polyprotein F;1-112/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                / Match 89.5%;
Local Similarity 89.3%;
nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 89.3 Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
ö
                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                       N.Contains: core protein
C:Species: hepatitis C virus
A:Variety: genotype 5, N2
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C.Species: heparitis C virus
A;Variety: genotype 1, N10
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Molecule type: genomic RNA
A; Residues: 1-114 < VAN>
A; Residues: 1-114 < VAN>
A; Cross-references: EMBL: 229471; NID: 9443904; PIDN: CAA82609.1; PID: 9443905
A; Experimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Cross-references: EMB::229445; NID:g443852; PIDN:CAA82583.1; PID:g443853
A; Experimental source: genotype 1, N10
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <MAT>
                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lipa
                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by
A;Reference number: S41341
A;Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 by
                                                                         Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 115;
                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R;van Doorn, L.J.: Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A;Description: Analysis of hepatitis C virus genotypes 1 to A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 137; DB 2; L
Pred. No. 1.7e-11;
1; Mismatches 2;
                                                                    Score 137; DB 2; I
Pred. No. 1.6e-11;
1; Mismatches 2;
C;Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 137; DB 2;
Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 1.66
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Mismatches
                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89.58;
                                                                      89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                    Query Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
```

us-09-491-146a-33.rpr

```
A)Cross-references: EMBL:229457
A)Experimental source: genotype 2, N3
C)Superfamily: hepatitis C virus genome polyprotein
C)Superfamily: hepatitis C virus genome polyprotein
C)Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <MAT>
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: S41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             completed: A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Search cor
Job time
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                    Nycontains: core protein.
C;Species: hepatitis C virus
A;Variety: genotype 1, N2
A;Variety: genotype 1, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41344
A;Description: A1344
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
A;Reference number: S41341
A;Accession: S41344
A;Accession: Accession: Ac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N8
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N;Contains: core protein
N;Contains: hepatitis C virus
A; Variety: genotype 2, N3
C; Date: 19-May 1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Cross-references: EMBL: 229453; NID:9443868; PIDN:CAA82591.1; PID:9443869 A; Experimental source: genotype 1, N8 C; Superfamily: hepatitis C virus genome polyprotein C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Cross-references: EMBL: 229447; NID: 9443856; PIDN: CAA82585.1; PID: 9443857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                     genome polyprotein - hepatitis C virus (genotype 1, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C,Accession: S41354
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R:van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Smbmitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Accession: $41350
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 115,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 137; DB 2; L
Pred. No. 1.7e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.5%; Score 137; DB 2; L
llarity 89.3%; Pred, No. 1.7e-11;
Conservative 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A: Experimental source: genotype 1, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKPQRKTKRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           uery Match 89.5%;
est Local Similarity 89.3%;
latches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA A; Residues: 1-115 < VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S41350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                          N; Contains: core protein
C; Species: hepatitis c virus
A; Variety: genotype 1, N3
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: EMBL:229448; NID:9443858; PIDN:CAA82586.1; PID:9443859
                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                               C,Accession: S41345
C,Accession: S41345
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa.
A; Reference number: S41341
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö

    hepatitis C virus (genotype 1, N3) (fragment)

       Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 137; DB 2; Length 115;
Pred. No. 1.7e-11;
1; Mismatches 2; Indels
    Score 137; DB 2; I
Pred. No. 1.7e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A: Experimental source: genotype 1, N3 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                      1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7, 2003, 11:21:49
    89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88.38;
Ouery Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  August
5 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
```

7

```
August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                   127863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                         tal number of hits satisfying chosen parameters:
                                                                                                                                                                                    US-09-491-146A-33
153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                       127863 seqs, 47026705 residues
                                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                      Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                             Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. SwissProt\_41:\*

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

.ศ.กินแบท DB seq length: 0 Maximum DB seq length: 2000000000

	000269 h genome po P27959 hepatitis c 001404 hepatitis c 001401 hepatitis c 201404 hepatitis c P27966 hepatitis c P27961 hepatitis c P26651 h genome po P26663 h genome po P26664 h genome po P28276 herpes simp P0958 homo sapien OB0341 murine corropistation of P3416 murine corropistation popular murine corropistation popular murine popular popu
SUMMARIES ID	POLG HCVUT POLG POND P POLG
DB	;
Length	3010 520 520 520 520 3013 3010 3010 3010 3010 3011 512 512 68 68 68 68 53 73 73 74 74 74 74 74 74 74 74 74 74 74 74 74
% Query Match	0.000000000000000000000000000000000000
Score	1337 1337 1337 1337 1337 1337 1337 1337
Result No.	

P40603 brassica na P03417 murine coro	015530 homo sapien 09z2a0 mus musculu	O55173 rattus norv Q03717 mus musculu	028824 bos taurus 099569 homo sapien	P35824 bacillus ci O15746 homo sapien	Q14684 homo sapien P26786 saccharomyc
APG_BRANA NCAP_CVMJH	PDPK_HUMAN PDPK_MOUSE	PDPK_RAT KCB1_MOUSE	KMLS_BOVIN PKP4_HUMAN	SLAP_BACCI KMLS_HUMAN	Y179_HUMAN RS7A_YEAST
					пп
449	556 559	559 857	1176	1616 1914	740 189
30.1	30.1	30.1	30.1 30.1	30.1	29.7
4 6 4 6	44.	4 P 4 6	4 4 6 4	4 4 6 6	45.5 45
34 35	35	3 6 3 7	4 4 0 1	4 4 3 2	4 4 4 5

## ALIGNMENTS

PECULT 1 POLG_HCVJT AC 000269; DD 1-APR-1993 DT 01-APR-1993 DE (GP68) (GP7) DE (GP68) (GP7); DE (GC 3.4.22.) RA MARAZAWA T., INTERCHED	HECVIT  POCAGE  Rel. 25, Created  POCAGE  Rel. 42, Last annotation upda  Genome polyprotein [Contains: Capsid prof  Genome polyprotein [Contains: Capsid prof  Genome polyprotein [Contains: Capsid prof  GENelope glycoprotein El (GP32) (GP35); BC  (GF6) (GP70) (NS1); Protein P7; Nonstruct  (GC 3.4.22); Protease/helicase NS3 (P70  (GC 3.4.21.98); Nonstructural protein NSA  NSSB (P65) (P70) (RNA-directed RNA Polyne  NSSB (P66) (P70) (RNA-directed RNA Polyne  NSSB (P66) (P70) (RNA-directed RNA Polyne  NCBI_TaxID=31642;  [1]  SEQUENCE FROM N.A.  MEDLINE-9229574; PubMed=1318627;  MEDLINE-229574; PubMed=1318627;  REDUINE-9229574; PubMed=1318627;  REDUINE-S29574; PubMed=1318627;  MEDLINE-229574; PubMed=1318627;  PROTEIN CAPACIVITY: Hydrolyala of four predicted individuals.;  Virus Res 23:39-53(1992).	STANDARD;  11. 25, Carl. 25, Lad.  12. 42, Lad.  13. 42, Lad.  14. 12, Lad.  15. 12, Lad.  16. 12, Lad.  17. 12, Lad.  18. 18. 18. 18. 18. 18. 18. 18. 18. 18.	STANDARD: PRT; 3010 AA.  (Rel. 25, Created) (Rel. 42, Last annotation update) (Rel. 42, Last annotation Nature (Rel. 42, Rel. 42, Rel. 1973) (Rel. 43, Nonstructural (RNA-directed RNA polymerase (Isolate HC-JT) (HCV). (Isolate HC-JT) (Isolate HC-JT) (HCV). (Isolate HC-JT) (I	uence otatio Capsid (GP3) (GP3	update) update) update) update) updates protein structur structur n NS4A ( n NS4A ( n NS2B, n	HEWAT STANDARD: PRT; 3010 AA.  900.0016;  10.1ARF-1993 (Rel. 25, Created)  10.1ARF-1993 (Rel. 25, Created)  10.1ARF-1993 (Rel. 26, Last sequence update)  10.1ARF-1993 (Rel. 26, Last sequence update)  10.1ARF-1993 (Rel. 27, Last annotation update)  660.000 (10.1ARF-1993) (Rel. 27, Last annotation update)  10.1ARF-1993 (Rel. 27, Last annotation update)  660.000 (10.1ARF-1993)  660.000 (10.1ARF-1
--	--	--	--	--	---	--

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92230232; PubMed=1314459;
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Trull-length sequence of a hepatitis C virus genome having poor ponotypes.;
Toplogy to reported isolates: comparative study of four distinct genotypes.;
Virology 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2B, NS2B, NS4B AND NS4B ARE HYDROPHOSIC. SUGGESTING A POSSIBLE MEMBRARE RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPORTOTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                            01-AŭG-1992 (Rel. 23, Created)
01-AŭG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last (GP32) (GP35); Envelope glycoprotein E2
16-OCT-2001 (Rel. 40, Repartitis C virus (isolate HC-12) (HCV).
17-OCT-2001 (Rel. 40, Rel. 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CAPSID PROFEIN C (POTENTIAL).
MATRIX PROFEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                       513 AA
                                              32
             Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPRO02522; HCV_capid.
InterPro; IPR002521; HCV_capid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_cenv.
InterPro; IPR002531; HCV_cenv.
InterPro; IPR002531; HCV_msl.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_env; 1.
ProDom; PF01550; HCV_msl; 1.
ProDom; PD16662; HCV_NSl; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10074; BAA00968.1; -.
                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                    POLG_HCVJ2
P27959;
                                           ß
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
CHAIN
CHAIN
TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                     RESULT 2
POLG_HCVJ2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                  CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
NAJOR ENVELOPE PROTEIN E (POTENTIAL).
NON-STRUCTURAL PROTEIN NSJ. (POTENTIAL).
PROTEASE/HELICASE NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.

CHARGE RELAY SYSTEM (BY SIMILARITY).

ATP (POTENTIAL).

DECH BOX.

N-LINKED (GLCNAC...) (POTENTIAL).

N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 1; Length 3010;
Pred. No. 2e-11;
1; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94A1C77435D642BB CRC64;
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002511; HCV_core.
InterPro; IPR002511; HCV_core.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR001218; HCV_NS3.
InterPro; IPR00140; HCV_NS3.
InterPro; IPR00145; HCV_NS4.
InterPro; IPR00145; HCV_NS4b.
InterPro; IPR00166; HCV_NS4b.
InterPro; IPR002166; HCV_NS4b.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01542; HCV_COTE; 1. Pfam; PF01542; HCV_COTE; 1. Pfam; PF01559; HCV_LNS1; 1. Pfam; PF01560; HCV_LNS2; 1. Pfam; PF012006; HCV_LNS3; 1. Pfam; PF01001; HCV_LNS4s; 1. Pfam; PF01001; HCV_LNS4s; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00988; Viral_RdRP; 1. CMART, 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                        PF01543; HCV_capsid;
PF01542; HCV_core; 1.
PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     729
1006
1615
1862
2013
3010
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          623
645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

ö

Gaps

ö

(POTENTIAL)

rearestantes en la company de la company de

 $\sim$ 

ò 셤

```
520 5
520 AA;
                                        Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transmembrane;
INIT_MET 1
                                                                                                                                                                                POLG_HCVHK
Q01403;
    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                            RESULT
                                                                                                                                                                                       SP
                                                                                           ò
                                                                                                                    셤
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CELLULAR AMINOPEPTIDAE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                            MDE K., Inchauspe G., Fullsawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
I. SUBUNIT THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
ILPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND MRNA.
                                                                                                                                                                                        01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
60-OCT-2001 (Rel. 40, Last annotation update)
Envelope 91ycoprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein El (GP32) (GP35); Envelope 91ycoprotein E2
(GP68) (GP20) (NS1)] (Fragment).
                                                     Gaps
                                                                                                                                                                                                                                                                       Hepatítis C virus (isolate HCV-476) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nonstructural protein.
REMOVED FROM CAPSID PROTEIN C
                         Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
            Score 137; DB 1; Length ox. Pred. No. 4.3e-12;
 943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ?
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.)
                                                                                                                                                                      520 AA
                                                   1; Mismatches
                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10688; BAA01530.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_capsid.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01545; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                     MEDLINE-93019030; PubMed-1383400;
 55704 MW;
                         89.5%;
89.3%;
                                    Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115
191
191
388
383
196
209
234
234
418
424
449
513 AA;
                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                              NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVH4
Q01404;
01-JUL-1993 (
                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
SEQUENCE
                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                         RESULT 3
POLG_HCVH4
ID POLG_H
AC Q01404
DT 01-JUL
DT 16-OCT
DE Genome
DE GENOELO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
SO
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Envelope divcoprotein El (GP22) (GP35); Envelope divcoprotein E2 (GP68) (GP70) (NS1)] (Fragment) (Hepatitis C virus (isolate HCV-KF) (HCV). Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepativius.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BY THE
                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PIR; JQ1925; JQ1925. | InterPro; IPR002522; HCV_capsid. | InterPro; IPR002521; HCV_core. | InterPro; IPR002521; HCV_core. | InterPro; IPR002531; HCV_env. | InterPro; IPR002531; HCV_env. | InterPro; IPR002531; HCV_env. | InterPro; IPR002531; HCV_env; | InterPro; IPR002531; HCV_env; | InterPro; IPR01539; HCV_env; | InterPro; IPR01539; HCV_env; | InterPro; IPR01540; HCV_
                                                                                                                Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nonstructural protein.

REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                Indels
520
56499 MW; AAl35246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                                                            Score 137; DB 1;
Pred. No. 4.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            520 AA.
                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BY SIMILARITY
                                                                                                                                                                                                                                                    1 PKPORKTKRKAHRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                         N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                         89.5%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115
191
191
369
209
233
234
418
```

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.sib.ch/announce/or send an email to license@isb.sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Transmembrane; Nonstructural protein.

INIT_MET

INIT_MET

I REMOVED FROM CAPSID PROTEIN C BY THE REMOVED FROM CAPSID PROTEIN C FOTENTIAL).

CHAIN

116

191

MADOR ENVELOPE PROTEIN E (POTENTIAL).

CHAIN

192

383

NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

CHAIN

734

734

NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein (Contains: Capsid protein C (Core protein); Matrix
Protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Viruses, SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2B, NS4B AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL, D10075; BAA00969.1; -.
InterPro: IPR00252; HCV_copsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002521; HCV_core.
InterPro: IPR002531; HCV_env.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01560; HCV_MS1; 1.
Probom; PB186062; HCV_MS1; 1.
Probom; PR018604; HCV_MS1; 1.
Probom; PR018604; HCV_MS1; 1.
Probom; PR018604; HCV_MS1; 1.
                                                                                Length 520;
                                                                           Score 137; DB 1; Length 52
Pred. No. 4.4e-12;
1: Mismatches 2; Indels
                                                     56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                               737 AA.
                                                                                                        1; Mismatches
                                                                                                                                 1 PKPORKTKRKAHRRPODVKFPGGGQIVG 28
                                                                                                                                                 89.5%;
89.3%;
                                                                              Ouery Match
Best Local Similarity 89.33
Matches 25; Conservative
424
431
449
520
424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11112;
                                                                                                                                                                                               RESULT 5
POLG_HCVJ5
ID POLG_HCVJ5
AC P27960;
                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genotypes.
                                        NON_TER
SEQUENCE
 CARBOHYD
               CARBOHYD
                          CARBOHYD
 FF FF 8
                                                                                                                                 ŏ
                                                                                                                                                          셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NSI and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-92230232; PubMed-1314459; Okamoto K., Lizuka H., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a heptitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                          Length 737;
                                                                                                                                                                                                                                                                                                                                                                                      Score 137; DB 1; Length 73
Pred. No. 6.3e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                               81207 MW; 3AF699D82AD501B1 CRC64;
                 N-LINKED (GLCNAC.
                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   737 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_sor.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 89.3%;
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10077; BAA00971.1; -:
                                                                                                                                                                                                                                                                                                                                                                                          89.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                               737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
TRANSMEM
                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                               Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
```

```
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
66nome polyprotedal [Contains: Capsid protein C (Core protein [P22);
670-68 (GP70) (NS1); Protein P7; Nonstructural protein R2 (GP8) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS44 (P4); Nonstructural protein NS58 (P66) (P70) (RNN-directed RNA polymerase) (EC 2.7.7.48)].
NS58 (P66) (P70) (RNN-directed RNA polymerase) (EC 2.7.7.48)].
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ./EZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSZ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Ontshi E., Andoh T., Yoshida I., Okayama H.;
Structure and organization of the hepatitis C virus genome isolated from human carriers.;
                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                          · · ) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE-96235224; Pubmed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 737;
                                                                  Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                  81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                        CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                        N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 137; DB 1;
Pred. No. 6.3e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKTKRKAHRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NSI; 1.
ProDom; PD186062; HCV_NSI; 1.
Polyprotein; Glycoprotein; Coat protein Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                           89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVBK
                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                              TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
                                                                                                                 CHAIN
                                                                                                                                        CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
              14441°
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

"" NS5a.
"AU07095; RNA\_POL\_DS\_PS.
"M; PF01543; HCV\_caps1d; 1.
"PF01543; HCV\_caps1d; 1.
"PF01542; HCV\_caps1d; 1.
"PF01542; HCV\_caps1d; 1.
"PF01542; HCV\_caps1d; 1.
"PF01542; HCV\_caps; 1.
"PF01542; HCV\_caps; 1.
"PF01542; HCV\_caps; 1.
"PF01542; HCV\_caps; 1.

Pfam;

PF01006;

InterPro; IPR007095; InterPro; IPR007094;

InterPro; IPR000745; InterPro; IPR001490;

InterPro; IPR002868; InterPro; IPR002166;

HCV\_capsid. HCV\_core.

InterPro; IPR001410; InterPro; IPR002522; InterPro; IPR002521; InterPro; IPR002531;

InterPro; IPR002519; InterPro; IPR002518; InterPro; IPR004109;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X PEDLINE-98227846; PubMed-9568891;

X An Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.;

X An Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.;

Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4 peptide of EM strain hepatitis C rius: a 2.2-A resolution structure in a hexagonal crystal form.";

Protein Sci. 7:837-847(1998).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL END RSA REPLICATION.

-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                   the viral
                                                                                       Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N., Moomaw E.W., Adachi T., Hostomska Z.; Habuka R.W., The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site."; Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                    X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; Pubmed-8861916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1A1Q; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB; 1JXP; 14-JAN-98
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1QUV; 26-JUN-00
BOHM; 20-APR-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [RNA](N).
```

```
POLG_HCVJA
P26662;
                   TURN
STRAND
SEQUENCE
  TURN
             HELIX
                                                                                                  POLG_HCVJA
                                                                                            RESULT
                                                                ò
                                                                          요
                                                                                                            FFFFFS
                                                        CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN (POTENTIAL).
MAJOR BUTCHIN (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTESAS/HELICASE NS3 (POTENTIAL).
PROTEXE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF001506; HCV_NS5a; 1.
Probom; PD186062; HCV_NS1; 1.
SWARF: SW00487; DEXDC: 1.
PO19protein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Cat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                    REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                (POTENTIAL).
                                                                                                                                                                                                                           (POTENTIAL).
                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
BECH BOX.
                                                                                                                                                  N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                 369
11083
11107
11165
11107
1319
1319
209
234
250
305
                                                                                                                                                                                                                                645
2041
2077
2240
22529
2788
1035
1047
1063
1074
                                                                                                                                                                                                                                                                                                          085
                                                                                                                                                                                                                                                                                                    .081
                                                                                                                                                                                                                                1122
1135
1135
1139
1149
1168
1162
1163
1172
                                                                                                                 347
1083
1107
1165
1230
1316
                                              3D-structure.
INIT_MET
                                                                                                                                 ACT_SITE
NP_BIND
SITE
                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                   CARBOHYD
                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                           SARBOHYD
                                                                                                                             ACT_SITE
                                                                                                                                                        CARBOHYD
                                                                                                                                                              CARBOHYD
                                                                                                                 TRANSMEN
                                                                                                                                                                                                                                                                        HELIX
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                    STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                    TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                 STRAND
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                   STRAND
                                                               CHAIN
CHAIN
CHAIN
CHAIN
                                                                                          CHAIN
                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                           TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation-the European Bioinformatics institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP22) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Proteases/helicase NS3 (P70) (Hepacivirin)
EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructur
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-91192160; PubMed-1849488;
Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Cato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Chikoshi S., Shimotohon K.;
"Molecular structure of the Japanese hepatitis C viral genome.";
FEBS Lett. 280:325-328(1991).
-! FUNCTION: THE SMALL PROPEIRS NS2A, NS2B, NS4A AND NS4B ARE
-! FUNCTION: THE SMALL PROPEIRE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-! CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-! CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ø,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDILINE-91088550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Appanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 1; Length 3010;
Pred. No. 2.8e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                             327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                  Query Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISCUSSION OF SEQUENCE.
                                                                                                    1202
1204
1688
1187 118
1189 119
1198 120
1203 120
1680 168
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11116;
```

```
PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA) (N)
                                                                                                                                                    POLG_HCVH
P27958;
                       ٦
                                                                                                                  RESULT
                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NAM - DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Prodom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 137; DB 1; Length 3010;
Pred, No. 2.8e-11;
l; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-L
                                                                                                                                                                                                                                    LucarPro; IPR007095; Helicase_C.
LucarPro; IPR007095; RNA_pol_DS_PS.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01542; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                           HCV_capsid.
                                                                                                            IPR002519; HCV_env.
IPR002518; HCV_NSI.
IPR002518; HCV_NSI.
IPR004109; HCV_NS3.
IPR000145; HCV_NS4A.
IPR001490; HCV_NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1
Pfam; PF00998; Viral_RGRP; 1
Pr0D0m; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2788
v; 327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89.5%;
                                                                                                                                                                                                                                                                                                                                                                         HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4b;
                                                                                        InterPro; IPR002521; H
InterPro; IPR002519; H
InterPro; IPR002531; H
InterPro; IPR002518; H
                                                      InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
HSSP; P26663; 1JXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                    PF01539;
PF01560;
PF01538;
PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                             PF01006;
PF01001;
PF01506;
                                                                                                                                                                                 InterPro; InterPro; InterPro; InterPro;
                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

Gaps

ö

Indels

89.3%;

Conservative

Local Similarity les 25; Conserv

Best Loca Matches

```
A MEDLINE-98154321; PubMed-94943270;

R AMAZY CMTSTALLOGARMATH (2.C. ANGSTROMS) OF 1192-105/.

R MEDLINE-98154321; PubMed-94943270;

R Min J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,

A MUTCKO M.A., Lin C., Caron P.R.;

A MUTCH OF MIN A 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions along as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annocation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); NONSTRUCTURAL (P65); NONSTRU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-57331322; Pubmed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTECUXIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genomic structure of the human prototype strain H of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       virus: comparison with American and Japanese isolates.".
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657
PRT; 3011 AA
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11108;
```

us-09-491-146a-33.rsp

```
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1627 1627
1635 1636
1640 1652
3011 AA; 327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 1;
Pred. No. 2.8e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
        CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                    TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
STRAN
          SO 1
        (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Probom; PD186062; HCV_NS1; 1.
SMART; SM06487; DEXDC: 1.
PO1yprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
3D-structure. 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
CHARGE RELAY SYSTEM (ECHARGE RELAY SYSTEM (ECHARGE RELAY SYSTEM (ECHARGE RELAY SYSTEM (ECHARGE BOX.
N-LINKED (GLCNAC...)
entities requires a license agreement ( or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               nterPro; IPR007095; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                             HCV_capsid.
                                                                                                                                                                                                                                                                             Interpro; IPR002519; HCV_env.
Interpro; IPR002511; HCV_NS1.
Interpro; IPR002518; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
Interpro; IPR001490; HCV_NS4a.
Interpro; IPR001490; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01543; HCV_capsid; Pfam; PF01542; HCV_capsid; Pfam; PF01542; HCV_capsid; Pfam; PF01550; HCV_cnv; 1. Pfam; PF01560; HCV_NS2; 1. Pfam; PF01006; HCV_NS3; 1. Pfam; PF01000; HCV_NS4; 1. Pfam; PF01001; HCV_NS4p; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00998; Viral_RGRP; 1. ProDom; PD186062; HCV_NS1; 1. ProDom; PD186062; 1. ProDom; PD186062; 1. ProDom; PD186062; HCV_NS1; 1. ProDom; PD186062; PD186062; 1. ProDom; PD186062; PD186062; 1. ProDom; PD186062; PD186
                                                     EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB; 1HEI; 25-NOV-98. PDB; 1A1V; 16-FEB-99. PDB; 1A1K; 17-JUN-98.
                                                                                                                                                           MEROPS; 539,001; ...
MEROPS; 039,001; ...
TRANSFAC; 704155; ...
InterPro; IPR001410; D
InterPro; IPR002522; H
InterPro; IPR002521; H
                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                            interPro; IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
```

ö

Gaps

us-09-491-146a-33.rsp

```
POLG_HCVJ8
P26661;
                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                         ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                  CHAIN
                                                                                                                                                                         CHAIN
                                                                                                                                      CHAIN
                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                     MEDLINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Amchida A., Miyakawa Y., Mayumi M.;
Amcleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved of divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
C -I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -I- CATALYITC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl'
C -I- CATALYITC ACTIVITY: N nucleoside triphosphate - N diphosphate +
      01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 21, Last sequence update)
01-40G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Viruses; SENNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D00944; BAA00792.1;
PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR001490;
                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                      NCBI_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
26-nome polyprotein (Contains: Capsid protein (Core protein) (P22);
66-nome polyprotein E1 (GP22) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3 4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3 4.22.-); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P66); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]
Hepatitis C virus (isolate HC-38) (HCV).
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
Probom; PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 CELLULAR AMINOPEPTIDASE.
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                           CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2811 N-LINKED (91273BE9E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 137; DB 1;
Pred. No. 2.9e-11;
1; Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
nes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus.
NCBI_TaxID=11115;
```

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last annoctation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS48 (P77); Nonstructural protein NS58 (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]
Hepatitis C virus (isolate Taiwan) (HCV).
Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           peptide bonds in the viral
            RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                   (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /; DB 1; Length 3vs:
. 2.9e-11;
-^hea 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1A173E7E3381FD1A CRC64;
                                             (BY
(BY
(BY
                                           CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                            N-LINKED (GLCNAC
                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 137;
Pred. No. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    330177 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ch 89.5%;
l Similarity 89.3%;
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                       2038 203
2359 235
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVTW
P29846;
                                                           ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                   CARBOHYD
                         TRANSMEM
ACT_SITE
                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
W Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
T INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
T CHAIN 115 CARSID PROTEIN (POTENTIAL).
T CHAIN 192 383 MAJOR ENVELOPE PROTEIN B. (POTENTIAL).
T CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
T CHAIN 1619 PROTEASE, HOLICASE NSI (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
T CHAIN 1630 NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
T CHAIN 1630 NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
T CHAIN 1630 NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
                                                                                                   SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
SEQUENCE FROM N.A. MEDMEd=1314459; OKAMOROLON, Lizuka H., Tanaka T., OKAMOROLO H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR007095; RNA_pol_DS_PS.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01518; HCV_NS2; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF01001; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR004109;
InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002166;
InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
```

CARAMA A TELEFORMAN A TELEFORMA

ö

Gaps

Gaps

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-91172826; PubMed-1848704;

MEDLINE-91172826; PubMed-1848704;

A Choo O.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,

Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Bradley D.W., Kuo G., Houghton W.;

"Genetic organization and diversity of the hepatitis C virus.";

Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1951).

"Correction THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

"S3 AND NS5 MAY PLAY A ROLE IN THE VIRAL FNA RELICATION."

"CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

"CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24). Last annotation update)
66-1992 (Rel. 24). Last annotation update)
67-1992 (Rel. 42). Last annotation update)
68-1992 (GPGB) (GPJB) (GPBB) (GP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                             (POTENTIAL)
                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                (POTENTIAL)
                                         POTENTIAL
                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                Length 3010;
     N-LINKED (GLCNAC. ..) (POTENT N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                        Score 130; DB 1; L
Pred. No. 2.8e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 KPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           623
645
2041
2077
2240
2788
1; 327047
                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 88.9%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
NCBI_TaxID=11104;
                                                                                       540
556
576
623
645
2041
2047
2240
22529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCV1
P26664;
                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCV1
          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        å
                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROYEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL,
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR00252; HCV_copsid.
InterPro; IPR00252; HCV_core.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002531; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR007045; HCV_NS4b.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007095; RNA_POL_PSYIR.
'fam; PF01543; HCV_copsid; I.
fam; PF01543; HCV_copsid; I.
fam; PF01550; HCV_CNS1; I.
'm; PF01539; HCV_NS1; I.
'm; PF01539; HCV_NS2; I.
'm; PF01503; HCV_NS2; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                             EMBL; M84754; -; NOT_ANNOTATED_CDS
PIR; A40244; GNWVTW.
PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_caps1d.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
Probom; PD186062; HCV_NS1; 1.
                        AND MRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                        MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             110863
2014
2014
1083
11083
111007
11230
1233
203
4117
4117
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDILINE-92113549; PubMed-1662697;
MEDILINE-92113549; PubMed-1662697;
"Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2."
J. Gen. Virol. 72:3057-3075(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FURI_HUMAN STANDARD; PRT; 794 AA.
P09958: 014336;
01-MAR-1989 (Rel. 10, Created)
01-MAR-1990 (Rel. 14, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Early protein; Transcription regulation; Activator; DNA-binding.
SEQUENCE 512 AA; 54958 MW; 459651470A503BA7 CRC64;
                                                                                             ö
                                               Score 129; DB 1; Length 3011; Pred. No. 4e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 53; DB 1; Length 512;
Pred. No. 4.6;
3; Mismatches 13; Indels
                                                                                          Indels
        327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                        Herpes simplex virus (type 2 / strain HG52).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID=10315;
                                                                                                                                                                                                                                                                                                                 01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Transcriptional regulator IE63 (VMW63) (ICP27).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          160 PKPRRRVSRNAHNQGGRHPASARTDGPGATHG 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58
                                                                                                                                                                                                                                                                                512 AA
                                                                                          2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRKAH-----RRPQDVKFPGGGQIVG
                                                                                                                                                     1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10471; BAA01269.1; -. EMBL; Z86099; CAB06702.1; -. PIR; JQ1498; WMBEXA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 37.5%;
Matches 12; Conservative
                                               Match 84.3%;
Local Similarity 82.1%;
les 23; Conservative
                                                                                                                                                                                                                                                                              STANDARD;
          3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                          IE63_HSV2H
P28276:
        SEQUENCE
                                                 Query Match
                                                                                                                                                                                                                                   RESULT 14
1E63_HSV2H
1D65_HSV2H
DT 01-DEC-
DT 28-FEB-
DE Transcr
DG UL54.
OS HETPES
OC ALPUSE
OC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
FURI_HUMAN
                                                                                        Matches
                                                                         Best
                                                                                                                                                                    8
                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR WATS PROTEIN (POTENTIAL).
NAJOR PROTEIN (SOTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                            Interpro; IPR004109; HCV_NS3.
Interpro; IPR00045; HCV_NS4.
Interpro; IPR001490; HCV_NS4.
Interpro; IPR002068; HCV_NSSa.
Interpro; IPR001166; HCV_RGRP.
Interpro; IPR001095; Helicase_C.
Interpro; IPR007095; RNA_POL_DS_PS.
Interpro; IPR007095; RNA_POL_DS_PS.
                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_corv.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR002511; HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01543; HCV_cap814] ...
Pfam; PF01543; HCV_cap814] ...
Pfam; PF01549; HCV_core; 1.
Pfam; PF01550; HCV_NS1; 1.
Pfam; PF01550; HCV_NS2; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF001001; HCV_NS4b; 1.
Pfam; PF00271; helicass_C; 1.
Pfam; PF00271; helicass_C; 1.
Pfam; PF00271; helicass_C; 1.
Pfam; PF00271; helicass_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    645
2041
2077
2240
2364
2789
PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                     InterPro; IPR001410;
InterPro; IPR002522;
                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 T_SITE_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
```

ä

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -.

Genew; HGNC:8568; FURIN.

MIM; 136950;

PIR; A39552; KXHUF. HSSP; Q99405; 1MPT. MEROPS; S08.071; -.

```
Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.; "Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
MEDLINE-94291619; PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilisin-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "cDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                                                 Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                          van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D., Dorssers L.C.J., van de Ven W.J.M.; "Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2."; Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
enzyme) (PACE) (Dibasic processing enzyme).
PACE OR FUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [5]
PROCESSING.
MEDLINE-92332543; PubMed-162922;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-91321735; PubMed-1713771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 296-794 FROM N.A. MEDLINE-87053858; PubMed=3023061;
                                                                                                                                                                                                                                                                                                                                     MEDLINE-90175002; PubMed-2408021;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                        NCBI_TaxID=9606
                                                                                                                                                                                                                                                                                                           TISSUE-Blood;
                  cleaving FURIN OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Brake A.
                  REPRETARY REPRET
```

```
ö
                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                       Score 53; DB 1; Length 794;
Pred. No. 7.4;
Wismatches 5; Indels
                                                                                                                                                                                                      CELL SURFACE SIGNAL,
TRANS GOLGI NETWORK SIGNAL.
10C44DD5892EF85D CRC64;
                                                                                                                                                                                                                                            :|:||| :: | | |||
102 KRRTKRDVYQEPTDPKFP 119
                                                                                                                                                                                                                                        21
                                                                                                                                                                                                               86678 MW;
                                                                                                                                                                                                                       34.6%;
50.0%;
                                                                                                                                                                                                                                        4 QRKTKRKAHRRPQDVKFP
                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                     553
75
107
500
762
                                                                                                                                                                                                                          Local Similarity
les 9; Conserv
                                                                                                                                                                                                               794 AA;
                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                              SITE
SITE
SITE
SITE
QQ
                                                                                                                                                                                                                                        ö
```

N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED

N-LINKED (GLCNAC...) (POTENTIAL)
CLEAVAGE (SECOND AUTO-).
CLEAVAGE (FIRST AUTO-).
CELL ATTACHMENT SITE (POTENTIAL).

(BY SIMILARITY)
(BY SIMILARITY)
(BY SIMILARITY)

CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM RELAY

POTENTIAL

24 794 794 705 738 738 153 194 368

FRANSMEM ACT\_SITE DOMAIN

CYS-RICH

POTENTIAL.

ACT\_SITE ACT\_SITE DISULFID DISULFID CARBOHYD CARBOHYD

protease; Transmembrane; Glycoprotein; Signal;

POTENTIAL.

Zymogen; Calcium

GO; GO:0005794; C:Golgi apparatus; TAS.
GO; GO:0004276; F:furin activity; TAS.
GO; GO:0007267; F:cell-cell signaling; TAS.
GO; GO:0007268; F:proteclysis and peptidolysis; TAS.
InterPro; IPR006312; Furin\_repeat.
InterPro; IPR00884; P\_domain.
InterPro; IPR000209; Peptidase\_88
Pfam; PF01083; P\_proprotein; PARTIAL.
PRINTS; PR00723; SUBTILISIN.
ProDom; PD00773; SUBTILISIN; 1.

7, 2003, 11:20:06 Search completed: August Job time: 5.90909 secs

-!- SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE EMDOPLASMIC RETICULM (ER).
SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN MOD THE ACTIVATION OF FURIN.
-!- TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.
-!- DOMAIN: CONTAINS A CYTOPLASMIC DOMAIN RESPONSIBLE FOR ITS TGN LOCALIZATION AND RECYCLING FROM THE CELL SURFACE.
-!- PTW: THE PROPEPTIDE IS AUTOCAPALYTICALLY REMOVED THROUGH AN INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM (ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD LEAD

++

,8888888888888888888888888888

-

COFACTOR: CALCION-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED

respective precursors.

PROPEPTIDE

<u>:</u>

This SWISS-PROT entry is copyright. It is produced through a collaboration

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY -!- SIMILARITY: Contains 1 homo B/P domain.

```
RESULT 1
                                                                                                         August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                   830525
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                             tal number of hits satisfying chosen parameters:
                                                                                                                                                                                       US-09-491-146A-33
153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                           830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_nammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_mhc:*
sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_plant:*
                                                                                                                                                                                       Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database :
                                                                                                                                                                                                                                                                                                                           Searched:
                                                                                                               Run on:
```

	c	ט פ	U	U	U	U	O	υ	O	U	O	O	O	Ų	O	c
t.ion	hepatitis	hepatitis	hepatitis	hepatitis	hepatitis		hepatitis	hepatitis	hepatitis	hepatitis	hepatitis	hepatitis				
Descrip	P90364	081831	Q9pxn3	068575	091au0	081989	081541	068311	068306	068309	Q8jyr9	081yr8	Q81yr5	Q8jyr7	Q8jyr6	068712
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1															
1D	P90364	081831	Q9PXN3	068575	<b>091AU0</b>	081989	081541	068311	908890	068309	Q8JYR9	Q8JYR8	Q8JYR5	Q8JYR7	OBJYR6	068712
DB	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
Length	105	106	125	191	3010	3010	3010	45	45	46	09	61	61	61	62	74
% Ouery Match	90.2	90.3	90.5	90.2	90.2	90.2	90.2	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5	89.5
Score	138	138	138	138	138	138	138	137	137	137	137	137	137	. 137	137	137
Result No.		8	۳,	4	Ŋ	9	7	80	6	10	11	12	13	14	15	16
	% Query Score Match Length DB	Score Match Length DB ID Descript 138 90.2 105 12 P90364	Guery Score Match Length DB ID 138 90.2 105 12 P90364 138 90.2 106 12 Q81831 081831	Score Match Length DB ID  138 90.2 105 12 P90364 P90364 P90364 P90364 138 90.2 105 12 Q91831 Q91831 Q91831	Acore Match Length DB ID  Score Match Length DB ID  138 90.2 105 12 P90364  138 90.2 105 12 Q9PXN3  138 90.2 125 12 Q9PXN3  138 90.2 191 12 Q88575  Q88575	Score Match Length DB ID Descript  136 90.2 105 12 P90364 P90364 138 90.2 105 12 O91831 O91831 138 90.2 125 12 O9PXN3 O9PXN3 138 90.2 3010 12 O68575 Q68575 138 90.2 3010 12 O31AUO O91AUO	Score Match Length DB ID Descript  138 90.2 105 12 P90364 P90364 138 90.2 105 12 Q81831 Q91831 138 90.2 125 12 Q9PXN3 Q9PXN3 138 90.2 191 12 Q68575 Q68575 138 90.2 3010 12 Q81989 Q81989	Score Match Length DB ID  138 90.2 105 12 P90364 138 90.2 105 12 Q91831 138 90.2 125 12 Q9PKN3 138 90.2 310 12 Q61AU0 138 90.2 310 12 Q81989 138 90.2 3010 12 Q81989 138 90.2 3010 12 Q81989	Ouery Score Match Length DB ID  138 90.2 105 12 P90364 138 90.2 106 12 QB1831 138 90.2 125 12 Q9FKN3 138 90.2 191 12 Q68575 138 90.2 3010 12 Q91AU0 138 90.2 3010 12 Q81989 138 90.2 3010 12 Q81989 138 90.2 45 12 Q68311 137 89.5 45 12 Q68311	Score Match Length DB ID Descript  138 90.2 105 12 P90364 P90364 138 90.2 105 12 O9PXN3 138 90.2 191 12 O9PXN3 138 90.2 3010 12 O9PXN3 138 90.2 3010 12 O9PXN3 138 90.2 3010 12 O9PXN3 138 90.2 45 12 O68361 137 89.5 45 12 O68361	Score Match Length DB ID  138 90.2 105 12 P90364 138 90.2 105 12 O9PXN3 138 90.2 125 12 O9PXN3 138 90.2 121 12 O6B575 138 90.2 3010 12 O9LN00 138 90.2 3010 12 OBL989 138 90.2 3010 12 OBL989 138 90.2 45 12 O6B311 137 89.5 45 12 O6B310 O6B311 137 89.5 46 12 O6B309 O6B311 137 89.5 46 12 O6B309	Ouery  Score Match Length DB ID  138 90.2 105 12 P99364 138 90.2 106 12 QB1831 138 90.2 125 12 Q9FKN3 138 90.2 125 12 Q9FKN3 138 90.2 3010 12 Q6575 138 90.2 3010 12 Q81589 138 90.2 3010 12 Q81589 138 90.2 45 12 Q68316 137 89.5 45 12 Q68316 137 89.5 60 12 Q81789 Q683109 Q683109 137 89.5 60 12 Q81789	Score Match Length DB ID Descript  138 90.2 105 12 P90364 P90364 138 90.2 105 12 OBH831 OBPXN3 138 90.2 125 12 OBPXN3 138 90.2 3010 12 OBH849 OBB575 138 90.2 3010 12 OBH541 OBH899 137 89.5 45 12 OG8301 OG8301 137 89.5 45 12 OG8309 OG8309 137 89.5 46 12 OG8309 OG8309 137 89.5 61 12 OSB3789 OG8378	Score Match Length DB ID  138 90.2 105 12 P90364 138 90.2 105 12 O9PXN3 138 90.2 125 12 O9PXN3 138 90.2 121 2 O9PXN3 138 90.2 3010 12 O9PXN3 138 90.2 3010 12 O9PXN3 138 90.2 3010 12 O91AU0 138 90.2 3010 12 O91AU0 138 90.2 45 12 O68311 137 89.5 45 12 O68310 137 89.5 46 12 O68310 137 89.5 60 12 O8B3YR9 137 89.5 61 12 O8B3YR9	Ouery Score Match Length DB ID  138 90.2 105 12 P99364 138 90.2 106 12 QB1831 138 90.2 105 12 QPKN3 138 90.2 125 12 QPKN3 138 90.2 3010 12 QB575 138 90.2 3010 12 QB1899 138 90.2 3010 12 QB1899 138 90.2 3010 12 QB1989 138 90.2 3010 12 QB1989 137 89.5 45 12 Q68316 137 89.5 45 12 Q68316 137 89.5 60 12 QB3YR8 137 89.5 61 12 QB3YR8 138 139 138 139 138 138 138 138 138 138 138 138 138 138	Score Match Length DB ID Descript    136   90.2   105   12   12   12   13   14   14   14   14   14   14   14

			۲,		۲,	٤,		rı		۲,	۲,		٠.	,,	*	,,	٠.	٠,	٠,	•	o		21		۲,	•	O	0			
Ø	o)	Ø	Ø	Ø	ιņ	o)	ø	Ø	ø	ø	Ø	Ø	co.	Ø	Ø	o)	Ø	Ø	co.	Ø	G	Ø	S	ø	co	ø	S	Ø			
111	111	iti	iti	iti	1t1	iti	1t1	1t.1	1t1	iti	iti	1t1	111	11	1t1	111	11.	iti	111	iti		titi	iti	111	iti	끍	iti	Ħ			
at	at	bat	bat	at	bat	bat	at	bat	at	at	at	at	at	at	at	at	bat	Sat	at	at	bat	at	at		at	Dat	at	bat			
her	her	he	hepatiti	hei	hei	he	her	hei	hei	her.	hei	he	hei	ь	hei	her	her	her	her	hei	hepatit	hepat	hepat	hepat	hepat	hepai	hepa	he			
82	9	79	83	90	84	07	85	s0	sı	64	61	20	24	57	22	88	26	90	64	55	9	23	28	63	61	22	54	29			
989	587	586	586	587	586	587	586	33,7	33.	312	3qt	Зďр	Здр	ðqt	gb	дb	3qt	Зąр	Зqр	ggp	Зgр	39p	3g.t	3 <u>q</u> p	Зqр	3qt	ğ	08qp59			
ŏ	ŏ	ŏ	ŏ	ă	ă	ŏ	ă	õ	õ	õ	ŏ	õ	õ	ö	õ	ŏ	ŏ	õ	ö	õ	õ	ö	ö	õ	õ	ö	ö	õ			
																													U	3	
																					•								STANKENTO	5	
	_	_	_			_		_				_		_		_		_			_	_		_		٠.	_	_	2		
682	708	679	683	706	684	707	685	ΥSC	YS1	264	T61	P70	P54	<b>T57</b>	P52	P86	156	<b>P9</b> 0	P64	P55	) 9d	P57	158	P63	P61	T52	T54	P59	1.1	1	
968682	968	968	998	268	998	268	998	287	287	281	290	88	88	8	280	380	280	280	280	380	980	980	96	280	88	8	8	280	4	•	
12	7	12	12	12	12	12	17	7	12	12	12	12	12	12	12	12	7	12	12	17	12	12	12	17	12	12	12	12			
7.4	74	74	74	74	74	74	74	9/	8	83	00	00	8	8	00	8	0	00	00	8	00	00	0	00	00	8	00	8			
•	•	•	•	•	•	•	•	•	•	_	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä	Ä			
9.5	•	•	•	٠	•	•	•	•		•	•		•	•	•	•	•	•	•		•	•	•				•	•			
æ	æ	æ	œ	æ	æ	æ	æ	æ	æ	æ	æ	œ	æ	æ	œ	80	œ	80	œ	œ	80	æ	80	8	80	æ	æ	ã			
7	7	7	7	7	7	7	_	7	7	_	7	7	_	~	7	7	7	_	_	7	7	_	7	7	7	7	7	7			
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13			
7	œ	o,	0	_	Ċ,	e	4	'n	و	7	æ	σ.	o	4	Ŋ	٣	4	'n.	ب	7	φ,	σ	0	_	N	3	4	'n			Ę.
-1		-	N	C	CI	~	N	CI	CI	CI	CI	CI .	m	m	m	m	m	m	m	m	m	m	4	4	4	4	4	4			RESULT
																															7

```
7
                                     Score 138; DB 12;
Pred. No. 1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 138; DB 12;
Pred. No. 1.6e-12;
                                                              le-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                       191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3010 AA
                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Mismatches
                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                        FILLIFIEL : FIRTIFIEL | FILLIFIEL | FILLIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; US5284; AAB00215.1; -.
InterPro; IPR002522; HCV_caps1d.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_core.
Pfam; PF01542; HCV_core; 1.
                               90.2%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 90.2%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
                                                                 Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q91AU0;
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                       068575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091AU0
                                                                                                                                                                                                                                                                                                            RESULT
Q68575
                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPART REPARTMENT OF THE PROPERTY OF THE PROPE
                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEGORASID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEGORASID IS A COMPLEX OF EMBL; D49465; BAAOBA39:1; -.
InterPro; IPR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-94351179; PubMed-7520922;
MREDLINE-94351179; PubMed-7520922;
MREDLINE-94351179; PubMed-7520922;
MREDLINE-94351179; PubMed-7520922;
J. Heparoll. 20:623-629(1994).
-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).

INTERPRO; IPRO02522; HCV_capsid;
Pfam; PF01543; HCV_capsid; 1.
                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
E2/NS1 protein (Genome polyprotein) (Fragment).
Hepartitis C virus type 2.
Hepartitis C virus type 2.
Hepartitis C virus type 2.
Hepartitis C virus type 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ohno T., Mizokami M.; **Determination of nine genotypes of hepatitis C virus using PCR
                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 12; Length 106; Pred. No. 8.8e-13;
   Length 105;
                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
(Isolate USA8) genomic RNA (Genome polyprotein) (Fragment).
                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             106 106
106 AA; 12001 MW; 25D0D5414B3EA9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        125 AA; 14220 MW; D945CB60661797E3 CRC64;
Score 138; DB 12;
Pred. No. 8.7e-13;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                     106 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. 8.8.
1; Mismatches
                                                                                                                      1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
   90.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 90.2%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
   Query Match 90.29
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=40271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
SEQUENCE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-USA8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                  081831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9PXN3
                                                                                                                                                                                                                                                                         RESULT 2
Q81831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                           SCORPTPROS
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-21440119; PubMed-11556407; Lim A.C., Lim S.G., Tan Y.H.; Lim S.P., Khu Y.L., Tay A., Ting A.E., Lim S.G., Tan Y.H.; Indentification and molecular characterization of the complete genome of a Singapore isolate of hepatitis C virus: sequence comparison with other strains and phylogenetic analysis."; Virus Genes 23:89-95(2001).
                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 191;
Length 125
                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20840 MW; 4AAE63444D8329E2 CRC64;
                                                                                                                                                                                                                           Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
```

Fri ..ug

```
single Japanese carrier in Nagasaki prefecture and genome analysis of E1 and E2/NS1 envelope glycoprotein regions.";
Jpn. J. Trop. Med. Hyg. 22:16941.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL. D63857; BAA09919.1; -.
HSSP; P26663; JJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.; "Molecular cloning of hepatitis C virus genome from a single Japanese Carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydrolase, Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 327503 MW; C7BDB38169D6E3CF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 12
Pred. No. 3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_POl_DS_PS
InterPro; IPR007094; RNA_POl_PSvir
Pfam, PP01543; HCV_capsid; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STKAIN-JI;
MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                       HCV_core.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS4b.
HCV_NS5a.
HCV_NS5a.
HCV_NS5a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01500; MCV_NS1; 1. Pfam; PF01500; MCV_NS1; 1. Pfam; PF012007; HCV_NS3; 1. Pfam; PF01000; HCV_NS4s; 1. Pfam; PF01001; HCV_NS4s; 1. Pfam; PF00271; hclicase_C; 1. Pfam; PF00271; hclicase_C; 1. ProDom; PF00999; Viral_RGRP; 1. ProDom; PD166062; HCV_NS1; 1. SWART; SW004487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90.2%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                            InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001490;
InterPro; IPR002868;
InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
nes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-JT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q81541
Q81541;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08154
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HAPPEN SERVICE SERVICE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane, SEQUENCE 3010 AA; 326793 MW; 3D89304314F9F795 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                     Lim S.P.;
Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-HCV-N; Zheng W.Z.; "Genotype identification of hepatitis c virus (HCV) isolated from "Genotype identification of hepatitis c virus (HCV) isolated from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
El and E2/NS1 envelope glycoprotein (Genome polyprotein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zheng W.-Y.Z.;
Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 12
Pred. No. 3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
||||||||||
5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                             EMBL, AF356827; AAL00900.1; --
InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR002522; HCV_copsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR004109; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00190; CYTOCHROME_C; 1. PROSITE; PS50507; RDRP_POSITIVE; 1. PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002166; HCV_RdRP.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01543; HCV_caps1d; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90.2%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; V1ral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 89.3 es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF02907; HCV_NS3; 1
PF01006; HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01001; HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
   SEQUENCE FROM N.A. STRAIN-HCV-S1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q81989
Q81989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
081989
AC 0819
AC 0819
DT 01-N
DT 01-
```

ö

Gaps

4

ö

Gaps

ö

```
STRAIN-HCV-BB5;
Songsivilai S., Kanistanon D., Kunkitti R.;
Identification and characterisation of Thai isolates of hepatitis C
       Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIVCORROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; U27749; AAA56056.1; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid.1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GINCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                   Score 137; DB 12; Length 45;
Pred. No. 5e-13;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                        CC527167096AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CC4C21ED236AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        068306 PRELIMINARY; PRT; 45 AA. 068306; 0.001-NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) 02-OCT-2016 (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 137; DB 12
Pred. No. 5e-13;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Last sequence upda
01-007-2002 (TrEMBLrel. 22, Last annotation updacer protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46 AA.
                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, U23744; AAA65051.1; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                   45
5015 MW;
                                                                                                                                                                                                                                                                                                     89.5%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45
5030 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                 45
45 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                        Polyprotein.
NON_TER 45
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                     Query Match
Best Local Simi
Matches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q68309;
Q68309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 9
Q68306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        268309

10 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 0000

20 000

20 000

20 000

20 000

20 000

20 000

20 000

20 00
            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polymerase; Transferase; Transmembrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glycoprotein; Nonstructural protein;
                 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HCV-BB37;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              05F0B2102CF9DD9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. 3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NS4B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
01-OCT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PUT.
                                                                                                                                                                                           InterPro; IPR002522; HCV_caps1d.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RNA-directed RNA
                                                                                                                                                                                                                                           IPR002519; HCV_env
IPR002531; HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              326564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4b; 1
HCV_NS5a; 1
                                                                                                                                                                    InterPro; IPR001410; DEAD.
Virus Res. 23:39-53(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 89.3 (atches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1006
1615
1862
2013
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2014 301
3010 AA;
                                                                                                                                             P26663; 1JXP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                068311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
068311
10 0683
AC 0683
AC 0683
DT 01-N
DT 01-
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
```

ö

Gaps

ö

Length 45; Indels

DB 12;

4

```
Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA (BY SIMILARITY).
    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                Score 137; DB 12; Length 61;
Pred. No. 6.9e-13;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                      61 61
61 AA; 6897 MW; EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (TrEMBLrel. 22, Created)
(TrEMBLrel. 22, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 12;
Pred. No. 6.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                             61 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                28
                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last seq
01-MAR-2003 (TrEMBLrel. 23, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                        InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AF50650; AAM33396.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                          89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
6897 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.5%;
                                                                                                                                                                                                                                                                       Query Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 89.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AA;
                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-RIG290;
                                                                                                                                                                                                                     Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-2002
01-0CT-2002
                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                          Q8JYR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8JYR
    õ
                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
        STRAIN-HCV-BB10;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis C
                                                                                 Ø
                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                              Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIZCORROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                 Length 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 137; DB 12; Length 60;
Pred. No. 6.8e-13;
                                                                                                                                                                                                    ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          B5691CBE8F26F2F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                             Score 137; DB 12;
Pred. No. 5.1e-13;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                 60 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 6.86
1; Mismatches
                                                                                                                                                                                                                                                                                   1 PKPQRKTKKKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AF506616; AAM33392.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                               InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment)
                                                                                                                                                                                                                             / Match 89.5%;
Local Similarity 89.3%;
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 60
60 AA; 6686 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.5%;
larity 89.3%;
Conservative
                                                                                                                                                                                        46 46
46 AA; 5129 MW;
                                                                                                                                EMBL; U23747; AAA65054.1;
                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
es 25; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-RIG306;
                                                                                                                                                                         Polyprotein.
NON_TER 46
SEQUENCE 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                               OBJYR9
OBJYR9;
                                                  virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OBJYR8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QBJYRB
                                                                                                                                                                                                                                                                                                                                                     RESULT 11
Q8JYR9
                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OBJYR8
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SETS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DT DT DT OS
```

Gaps

ö

Length 61; Indels

ö

Gaps

ö

K

ΒY

QF

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                             Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitteed (Apr. 202) to the EMBL/GenBank/DDBses.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

InterPro; IPR002522; HCV_capsid.
Pfan: PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMEL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                      Viruses; ssrNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genome polyprotein (Fragment).
Heparitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 137; DB 12; Length 61;
Pred. No. 6.9e-13;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 89.5%; Score 137; DB 12; Length 62; Best Local Similarity 89.3%; Pred. No. 7e-13; Matches 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 61 .
61 AA; 6869 WW; EC657F706E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EEEC656DC79E8F26 CRC64;
01-MAR-2003 (TrEMBLrel. 23, Last annotation update) Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPORKTKRKAHRRPODVKFPGGGQIVG 28
|||||||||||||||||||||||||||||||||||5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF506619; AAM33395.1; -...
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 62
62 AA; 7053 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                           [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                     NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                  STRAIN-RIG288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-RIG289;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER
SEQUENCE 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OBJYR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08JYR6
1D 08B
AC 09B
DT 01D
DT 01D
DT 01D
DC 06B
OC HE
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
```

Search completed: August 7, 2003, 11:19:03 Job time : 25.6364 secs

us-09-491-146a-33.rai

```
August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                          1: /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/laa/pcTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/laa/pcTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-921-887-33
US-08-921-887-28
US-08-921-887-28
US-08-635-886C-191
US-08-380-160-6
US-09-020-846-36
US-09-389-160-1
US-08-380-160-2
US-09-389-156-1
US-08-380-160-2
US-08-380-160-3
US-08-262-037-26
PCT-US93-08638-23
PCT-US93-08638-23
PCT-US95-13660-3
                                                                                                                                                                                                                                                     tal number of hits satisfying chosen parameters:
                                                                                                                                US-09-491-146A-33
153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                              328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
                                                     - protein search, using sw model
                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                 Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                               Issued_Patents_AA:*
                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length
                                                                                                                                                                                       BLOSUM62
                                                                                                                                                                                      Scoring table:
                                                                                                                                  Title:
Perfect score:
                                                   OM protein
                                                                                                                                                                                                                                                                                                                                                                             Database :
                                                                                                                                                            Sequence:
                                                                                                                                                                                                                             Searched:
                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ٠
ي
```

PP11 P111 P111 P111 P111 P111 P111 P11	THE SAN	ő
Applitable		
60, Appl 1, Appli 2, Appli 16, Appl 15, Appl 156, App 157, App 159, App 160, App 161, App 163, App 164, App 165, App 165, App	FOR MAKING	Gaps
	<del>v</del>	
Sequence Seq	F01	ő
sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce Sedneuce	RESULT 1  US-08-921-887-33  SEQUENCE 33, APPLICATION US/08921887  SEQUENCE 31, APPLICATION: PATENT NO. 6030771  GENERAL INVENTION: APPLICANT: KHUDYAROV, VURI E. APPLICANT: FIELDS, HOWARD A. TITLE OF INVENTION: MOSALC PROTEIN AND RESTRICTION TITLE OF INVENTION: MOSALC PROTEIN AND RESTRICTION TITLE OF INVENTION: MOSALC PROTEIN AND RESTRICTION TITLE OF INVENTION: MOSALC PROTEIN AND REPORT OF THE OF INVENTION: MOSALC PROTEIN AND RESERVENCE ADDRESS: CORRESPONDENCE ADDRESS: ATABLE OF INVENTION: MOSALC PROTEIN AND ALLA ALLA ALLA ALLA ALLA ALLA ALLA	Length 28; Indels
	N N O	Len
50 62 62 62 63 115 115 115 116 116 116 116 116 116 116	Floor #1.30	3; -15; 0;
US-08-836-075A-60 US-08-501-195-1 US-08-51-115-1 US-08-54-511-2 US-07-681-701-16 US-08-290-665A-155 US-08-290-665A-155 US-08-290-665A-156 US-08-290-665A-156 US-08-290-665A-169 US-08-290-665A-169 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-161 US-08-290-665A-164 US-08-290-665A-164 US-08-290-665A-164 US-08-290-665A-164	AND RESTED 1:, 37th	ore 153; DB ed. No. 5.8e Mismatches
90000000000000000000000000000000000000	ALIGNMENT TEIN AND SE ASSIST , N.W., 3 , N.W., 3 1,887 1,887 63-0380	153; No. natch
8-102222222222222	OS OS (10 )	Score Pred. Misn
2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2	ALIGNMEN  US/08921887  YURI E. WARD A. SALIC PROTEIN AND NDONUCLEASE ASSIS SS SS SS SS SS ASKEW, LLP ee Street, N.W., ee Street, N.W., release #1.0, Ver Release #1.0, Ver Release #1.0, Ver TON: ASKEW, LLP EST. 03063-0380 WASTION: 36,714 36,714 36,714 36,714 36,714 36,714 36,714 36,714 36,714 36,714 36,714 36,714 31: 31: 32: 33: 33: 33: 33: 34: 34: 34: 34: 34: 34	S.T.
	Application US/089219 30371 RWATION: KHUDYAKOV, YURI E. FIELDS, HOWARD A. NYENTION: MOSAIC PR. NYENTION: MOSAIC PR. SEQUENCES: 55 SEQUENCES: 55 SEQUENCES: 55 SANCE ADDRESS: USA 303-1769 SADABLE FORM: FPE: Floppy disk TPE: Floppy disk TONA TON NUMBER: US/08/99	., .O
	JS/08 AARD AARD DOONU DO	100.0%; 100.0%; 1ve
11000000000000000000000000000000000000	DYAKOV, ) DYAKOV, ) DYAKOV, ) DYAKOV, ) DYAKOV, ) DON: BOI DON: BO	100 larity 100 Conservative
	HIDDIARYOV, HUDDIARYOV, HUDDIARYOV, HIDELDS, HO NYTION: NYTION: NYTION: HIDELDS, HO NYTION: HIDELDS, HO HIDELS, HO HIDELS	.y IIVa
ល់កំណុកកំណុកកំណុកកំណុកកំណុកកំណុ	ON NOTICE OF CONTROL O	rit nse
O O O O O O O O O O O O O O O O O O O	13 APP11 APP11 APP11 KHU KHU KHU KHU KHU KHU LIDAL LID	11a; Coj
	T. 1  -921-887-33  uence 33, Applicuence 33, Applicant No. 6030771  APPLICANT: KHUD APPLICANT: KHUD APPLICANT: FIEL TITLE OF INVENTI WUMBER OF SEQUEN CORRESPONDENCE A ADDRESSE: JO STREET: 191 P CITY: A13013-17 COMPUTER: LBM TELEFOMM TELEFOMM TELEFOMM TELEFOMM TELEFOMM TELEFOMM TELEFOMM TYPE: MAIN CANTO TOPOLOGY: NO STRAMDEDMESS: LBM TYPE: MAIN CANTO COMPUTER: NO STRAMDEDMESS: LBM TYPE: MAIN CANTO COMPUTER: NO STRAMDEDMESS: LBM TYPE: MAIN CANTO COMPUTER: NO STRAMDEDMESS: COMPUTER: NO STRAMDE	h Similarity 28; Conser
	1.21-887-3 inice 33, inice 33, inice 33, inice 33, inice 33, inice 33, inice 34, inice 34, inice 34, inice 36, inice	tch al 2
·	TELECT APPLIANCE OF THE COMPORT OF T	y Matc} Local hes
ИИМИМИМИМИМИЧ <b>4444</b> В Ф О Н И М 4 И Ф C P O C P	RESULT 1  US-08-921-887-33  Sequence 33, Application US/08921887  Patent No. 6030771  GENERAL INFORMATION: PATENTE OF INVENTION: MOSALC PROTEIN F TITLE OF INVENTION: ENDONUCLEASE ASS NUMBER OF SEQUENCES: 55 CORRESPONDENCE ADDRESS: 55 COUNTRY: GLAR STREE: 191 Peachtree Street, N.W. CITY: Atlanta STRATE: GAN COUNTRY: USA COUNTRY: USA COUNTRY: USA STREET: 191 Peachtree Street, N.W. CITY: Atlanta STRATE: GAN COUNTRY: USA COUNTRY: USA COUNTRY: USA STREET: 191 PEACHTINE STREET COUNTRY: USA SOFTWARE: PATENTION DATA: APPLICATION NUMBER: US/08/921,887 FILING DATE: CLASSIFICATION UNDER: 36,714 REFERENCE/DOCKET NUMBER: 03063-038 TELECHOMON: AO4-818-3700 INFORMATION FOR SEQ ID NO: 33: SEQUENCE CHARACTERISTICS: LENGTH: 28 amino acids TYPE: Internal ONGGANISM: HEPATLIS VITUS USGANISM: HEPATLIS VITUS USGANISM: HEPATLIS VITUS	Query Match Best Local : Matches 2:
	O C C C C C C C C C C C C C C C C C C C	

Sequence 10, Appl Sequence 198, App Sequence 232, App Sequence 14, Appl Sequence 8, Appl Sequence 8, Appli Sequence 8, Appli Sequence 8, Appli Sequence 2, Appli

US-08-836-075A-10 US-08-635-886C-198 US-08-635-886C-232 US-08-836-075A-14

2, A 46, A

Sequence

US-08-836-075A-46

us-09-491-146a-33.rai

```
Sequence 191, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
APPLICANT: LEROUX-ROBLS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MARTENS, Geert
TITLE OF INVENTION: UMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: UMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2752-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 4; Length 450;
Pred. No. 1.2e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 90.2%; Score 138; DB 3; Length 28; Best Local Similarity 89.3%; Pred. No. 7.2e-13; Matches 25; Conservative 1; Mismatches 2: Indele
                                                                                                                                                             MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CURRENT APPLICATION NUMBER: US/08/635,886C CURRENT FILING DATE: 1996-04-25 PRIOR APPLICATION NUMBER: PCT/EP94/03555 PRIOR FILING DATE: 1994-10-28 PRIOR APPLICATION NUMBER: EP 93402718.6 PRIOR FILING DATE: 1993-11-04
                                                                                                                                                                                                                                                                                                                             CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REPERENCE/POCKET UMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHEICAL: NO ANTI-SPACE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PatentIn version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: hepatitis C virus
US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGIH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    internal
                                                                                         COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version
SEQ ID NO 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: not
                                                  Atlanta
                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOGALC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 25, Application US/08921887

Patent No. 603071

GENERAL INFORMATION

APPLICANT: KHUDYAKOV, YURI E.

APPLICANT: FIELDS, HOWARD A.

TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION

TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: USA
ZIP: 30103-1769
ZIP: 30103-1769
COMPUTER READBLE FORM:
MEDLUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPPRATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: PAtentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
FILING DATE:
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96.7%; Score 148; DB 3;
96.4%; Pred. No. 2.9e-14;
tive 0; Mismatches 1;
                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    03063-0380
                                                                                                                                                                                 Sequence 28, Application US/08921887
Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/CDCKET NUMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: not relevant TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-CENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 96.4<sup>3</sup>
Matches 27, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                        Atlanta
: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE:
TELEFAX: 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-921+887-25
                     ò
                                                                 а
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

ö

ö

Gaps

```
Score 137; DB 3; Length 34;
Pred. No. 1.2e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY. U.S.A.

ZIP: 2007-5109

COMPUTER READABLE FORM:
MEDIOW TYPE: FILIPPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846

FILING DATE: 09-FEB-1998
CLASSIFICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAMME: WEGIGE. HAROR
REGISTRATION NUMBER: 25,258
REFERENCE/POCKET NUMBER: 25,258
REGISTRATION NUMBER: 25,258
REFERENCE/POCKET NUMBER: 053466/0225
TELECOMMUNICATION NUMBER: 25,258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: YAWAGUCHI, Kenjiro
APPLICANT: KASHIMAKUMA, Tomiko
APPLICANT: CHIBA, Yukie
APPLICANT: YAGI, Shintaro
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAWA, Akira
TILE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FRAGMENT TYPE: N-terminal ORIGINAL SOURCE: CRANISM: Human Hepatitis C Virus US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 36, Application US/09020846
Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: 3000 K Street, N.W. CITY: Washington STATE: D.C. COUNTRY: U.S. A
REFERENCE/DOCKET NUMBER: WF
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 89.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (202) 672-5399
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                              single
                                                                                                                                                                                                                                          TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                     ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JS-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 6, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
PAPELICANT: JOLIVET, Michel
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                        APPLICANT: DELEYS, ROBERT APPLICANT: MARRIENS, GEART TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS (TITLE OF INVENTION: URMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS (CURRENTE OF INVENTION: VIRUS (UNDER: 2752-18)

FILE REFERENCE: 2752-18

FILE R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 4; Length 450;
Pred. No. 1.2e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                         5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-WAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                    Sequence 192, Application US/08635886C
Patent No. 655114
PAPERAL INFRMATION:
APPLICANT: LEROUX-ROELS, Geert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OLIFF & BERRIDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Berridge, William P. REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQ ID NOS: 286
SOFTWARE: PatentIn version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; TYPE: PRT; ORGANISM: hepatitis C virus US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 89.3%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET: P.O. Box 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                             -08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: V. COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 192
LENGTH: 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-380-160-6
                         ò
                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

Gaps

```
Score 137; DB 3; I
Pred. No. 1.6e-12;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ch 89.5%; Score 137; DB 4; Stailarity 89.3%; Pred. No. 1.6e-12; 25; Conservative 1; Mismatches 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
                  Query Match
Best Local Similarity 89.38;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                           RESULT 9
US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                           Sequence 2, Application US/08380160
| Patent No. 6235284
| GENERAL INFORMATION:
| APPLICANT: DALBON, Pascal | APPLICANT: JOLIVET, Michel | TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE | TITLE OF INVENTION: FOR DETECTING THE LATTER | TITLE OF INVENTION: FOR DETECTING THE LATTER | TITLE OF INVENTION: FOR DETECTING THE LATTER | ADDRESSE: OLIFF & BERRIDGE | ADDRESSE: OLIFF & BERRIDGE | STREET: P.O. BOX 19928
                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note" "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                  ö
                                                                                                              Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 22320
ZIP: 22320
MEDIUM TYPE: Floppy disk
COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURREMY APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                 Query Match 89.5%; Score 137; DB 4; L
Best Local Similarity 89.3%; Pred. No. 1.5e-12;
Matches 25; Conservative 1; Mismatches 2;
                                                                                                                                                                             1 PKPORKTKRKAHRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: Beridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION: 1.44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S
N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE
                                              ; MOLECULE TYF
US-09-020-846-36
                                                                                                                                                                                                                                                                           RESULT 8
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-380-160-2
                                                                                                                                                                               à
                                                                                                                                                                                                               q
```

```
Sequence 1, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
TITLE OF INVENTION: FOR DETECTING THE LATTER
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                              Sequence 1, Application US/09389756
Sequence 1, Application US/09389756
Sequence 1, Application US/09389756
Sequence 1, Application US/09389756
Sequence 1, Application Sequence 1, Application Sequence 1, Applicant: Joliver 1, Applicant: Dalony Pages 1, Applicant: Lacoux, Xavier Applicant: Lacoux, Xavier 1, Interpression Deference 1, Applicant 1, 103959
TITLE OF INVENTION: COMPOUND: NICOURE 1, 103959
CURRENT FILING DATE: 1999-09-07
BARLIER FILING DATE: 1999-09-07
BARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
LENGTH: 44
                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Hepatitis C virus
ORGANISM: Hepatitis C virus
FUBLICATION INFORMATION:
AUTHORS: Ogate, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
FITLE: Off Hepatitis Virus
COURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 1991
DATE: 1991
ACES: 1392-1396
DATE: 1991
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 22320
COMPUTER READABLE FORM:
```

```
Sequence 23, Application US/08083947

Patent No. 5639594

GENERAL INFORMATION:
Patent No. 7639594

GENERAL INFORMATION:
TITLE OF INVENTION: In Diagnosing and Detecting No. 5639594-1

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
CITY: Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                Score 137; DB 1; Length 61;
Pred. No. 2.2e-12;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 137; DB 1; Length 61; Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGENT INFORMATION:
  APPLICATION NUMBER: US/07/946,054 FILING DATE: 15-SEP-1992 CLASSIFICATION: 435
                                                                             NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFRENCE/DOCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
INFORMATION FOR SEQ ID NO: 9:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION: INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.5%;
                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (516)273-2828
                                                            ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Floppy disk
                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: US
ZIP: 11.788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE:
                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequent No. 5582968

Batent No. 5582968

GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
APPLICANT: Wang, Chang Y1
APPLICANT: Hosein, Barbara H
TITLE OF INVENTION: No. 5582968e1 Branched Hybrid and Cluster
TITLE OF INVENTION: No. 5582968 B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
CITY: Hamppauge
CITY: Hamppauge
CITY: Hamppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89.5%; Score 137; DB 3; Length 45;
89.3%; Pred. No. 1.6e-12;
11ve 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note- "N-terminal sequence of the
                                 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION UNMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                       CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: BETINGE, William P.
REGISTRATION NUMBER: 30,024
REFERNEX/POCKET NUMBER: WPB 28682
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FRACMENT TYPE: N-terminal ORIGINAL SOURCE: ORGANISM: Human Hepatitis C Virus STRAIN: H77
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 89.39
Matches 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                             45 amino acids
                                                                                                                                                                                                                                                                                                                                             TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: 1.45
OTHER INFORMATION: /
OTHER INFORMATION: P
OTHER INFORMATION: t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: Peptide
LOCATION: 1..45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: New York
COUNTRY: USA
ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
```

```
Sequence 9, Application PC/TUS9308638 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEX: 421792
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 89,3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Chang Y1 Wang and Barbara Hosein
TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
CORRESPONDENCE ADDRESS:
        Gaps
                                                                                                                                                                                    Sequence 3, Application US/08530550
Patent No. 5736321
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TILLE OF INVENTION: Peptides Effective for Diagnosis and
TILLE OF INVENTION: Detection of Hepatitis c Infection
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.5%; Score 137; DB 1; Length 61; 89.3%; Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: COMPUTER: COMPATION SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 26, Application US/08262037; Patent No. 5747239; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/POCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
CITY: Hauppauge
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        uery Match 89.5
est Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM: MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: peptide US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                    US-08-530-550-3
Matches
                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
APPLICANT: United Biomedical Inc.
TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
TITLE OF INVENTION: Non-B Hepatitis
NUMBER OF SEQUENCES:
ADDRESSE: UNITED BIOMEDICAL INC.
STREET: 25 Davids Drive
CITY: Hauppauge
STATE: New York
COUWRRY: USA
ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89.5%; Score 137; DB 1; Lv
89.3%; Pred. No. 2.2e-12;
11ve 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC CONDATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUBER: US/08/262,037
FILING DATE: A-JUBER: US/08/275
FILING DATE: 11-MAI-1991
APPLICATION NUMBER: US/651,735
FILING DATE: C-JUBER: US/58,799
FILING DATE: A-JUBER: US/58,799
FILING DATE: A-JUBER: US/58,799
FILING DATE: A-JUBER: US/58,799
FILING DATE: A-JUBER: US/58,799
FILING DATE: US-MARTION:
NAME: MAIA C. H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4043 (
TELECOMMUNICATION NUMBER: 1151-4043 (
TELEPHONE: 212-758-4801
FILING TATES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: PCT/US93/08638
```

```
us-09-491-146a-33.rai
```

0; Gaps

Search completed: August 7, 2003, 11:23:53 Job time : 10.6364 secs

```
August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1970) (1
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  153
1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                               OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-491-146A-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                          Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Sequence 42, Appl Sequence 44, Appl
US-09-878-281-42 US-09-878-281-44
11
169 169
89.5 89.5
137 137
14 15

89.5%; Score 137; DB 15; Length 44;

Query Match

Sequence 2, Appli Sequence 3, Appli Sequence 16, Appli Sequence 18, Appli Sequence 18, Appli Sequence 50, Appli Sequence 51, Appli Sequence 52, Appli Sequence 52, Appli Sequence 50, Appli Sequence 50, Appli Sequence 50, Appli Sequence 50, Appli Sequence 50, Appli Sequence 50, Appli Sequence 21, Appli Sequence 22, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 26, Appli Sequence 27, Appli Sequence 27, Appli Sequence 27, Appli Sequence 27, Appli Sequence 27, Appli	IGENIC AND IMMUNOGENIC G, PREVENTING AND PCT/FR98/00442 H Strain
0 US-09-929-955-2 0 US-10-104-966-3 1 US-09-194-949-3 0 US-09-185-138-76 0 US-09-851-138-76 0 US-09-851-138-18 1 US-09-899-046-50 1 US-09-899-046-50 1 US-09-899-046-54 1 US-09-899-046-54 1 US-09-899-046-54 1 US-09-878-281-52 1 US-09-878-281-54 1 US-09-978-808-50 1 US-09-978-808-50 1 US-09-995-860-50 1 US-09-995-860-50 1 US-09-995-860-50 1 US-09-995-860-50 1 US-09-952-75-40 1 US-09-952-75-40 1 US-09-952-75-40 1 US-09-952-955-11 1 US-09-952-955-11 1 US-09-952-955-11 1 US-09-923-955-11 1 US-09-923-955-11 1 US-09-923-955-11 1 US-09-923-955-11 1 US-09-923-955-11	ALIGNMENTS  TOTURAL PEPTIDE, ANT ND USES FOR DETECTION  TAGT, 677  B9,756  APPLICATION NUMBER: DATE: 1998-03-05  LIATION RATE Of the U.S.A.
182 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 1982 140 140 140 140 140 140 140 140 140 140	A11 III III SSC
30     30	SSULT 1 S-10-367-677-1 Sequence 1, Application US/1036767 Sequence 1, Application US/1036767 Publication No. US20030118604A1 GENERAL INFORMATION: APPLICANT: JOLIVET, MICHEL APPLICANT: LADALBON, PASCAL APPLICANT: LADALBON, ANTIGENIC STRITLE OF INVENTION: ANTIGENIC STRITLE OF INVENTION: ANTIGENIC STRITLE OF INVENTION: ANTIGENIC STRITLE OF INVENTION NUMBER: US/09/39 PRIOR FILING DATE: 1999-09-07 PRIOR FILING DATE: 1999-09-07 PRIOR FILING DATE: BARLIER FILING NUMBER OF SEQ ID NOS: 11 SSCOTUB NO 1 LENGTH: 44 TYPE: PRT ORGANISM: Hepatitis C virus PUBLICATION INFORMATION: AUTHORS: Ogata, N. et al. TITLE: Nuclectide Sequence and MATHORS: Ogata, N. et al. TITLE: Nuclectide Sequence and MATHORS: 094214, N. et al. TITLE: Nuclectide Sequence and MATHORS: 1991 PASCES: 1991 PREEEVANT RESIDUES: 2 TO 45 S-10-367-677-1
11111111111111111111111111111111111111	SULT 1 -10-367-67-1 Sequence 1, Applicat Sequence 1, Applicat Sequence 1, Applicat GENERAL INFORMATION: APPLICANT: JOLIVET, APPLICANT: LADANTE APPLICATION FILLE OF INVENTION: FILLE REFERENCE: 103 CURRENT FILLING DATE: PRIOR FILLING DATE: PRIOR PRILING DATE: PRIOR APPLICATION N PRIOR FILLING DATE: PRIOR PRILING DATE: PRIOR APPLICATION N PRIOR PRING DATE: PRIOR PRING DATE: PRIOR APPLICATION N PRIOR PRING DATE: PRIOR APPLICATION N PRIOR APPLICATION N PRIOR APPLICATION N SEG ID NO 1 LENGTH: 44 LENGTH: PRI LENGTH: PRI LENGTH: PROCOLUME: 88 PAGES: 3392-3396 DATE: 1991 RELEVANT RESIDUES: -10-367-677-1
11112222222222222222222222222222222222	RESULT 1 Sequence 1, Sequence 1, Publication GENERAL INFC APPLICANT: PRIOR FILIN NUMBER OF 5 SOFTWARE: PRIOR FILIN NUMBER OF 5 COTTUBE: NUC TITLE: NUC T

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 97;
                                                                                                                                                                                         ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 137; DB 10;
Pred. No. 2.7e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERALINFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3: Suite 701-E Columbia Square 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                   1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: PCT/GB93/00410 FILING DATE: 26-FEB-1993 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
                          60/092,339
   CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 60/092,
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOCTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 18
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89.5%;
                                                                                                                                                                                                             89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US 08 FILING DATE: 29-AUG-1994
                                                                                                                                TYPE: PRT
ORGANISM: Hepatitis C Virus
                                                                                                                                                                                                   Query Match
Best Local Similarity 89.38
Matches 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET: 555 13th SCITY: Washington STATE: D. C. COUNTRY: U. S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                        US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
FILE REFERENCE: 14114.0349U2
                                                                                                                                                                                                                                                   TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: BP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 74;
                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
 Pred. No. 1.2e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.5%; Score 137; DB 10; 89.3%; Pred. No. 2.1e-11;
                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                     1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                      4 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
                                                                                                                                                           US-09-851-138-10
Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                  AGENTS
                                                                                                                                                                                                                                                                                                                                                                                       STREET: P.O. BOX 4433
CITY: HOUSTON
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 89.59
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                 NUMBER OF SECUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: TEXAS
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 3
US-09-758-308-1
                                                                                                                                             RESULT 2
                                                                                    셤
                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: Dathogenic strain of the hepatitis C virus and TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ; CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT APPLICATION NUMBER: EP 00402225.7
PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARR: PATENTIN VET. 2.1
SEQ ID NO 78
LENGTH: 113
                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                              ·.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 0;
                                                                       Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
                                                                Score 137; DB 10;
Pred. No. 3.1e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 137; DB 10;
Pred. No. 3.2e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSPON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                          1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                 1 PKPORKTKRKAHRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; SEQUENCE DESCRIPTION: SEQ ID NO: 14: US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 08/836,075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 46, Application US/09851138 publication No. US20020183508Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                Sequence 78, Application US/09921397 Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: MAERTENS, GEERT STUYVER, LIEVEN
                                                                  Query Match 89.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-851-138-46
                                                                                                                                                                                                                                                                                            US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                          å
                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HEPATITIS C VIRUS GENOTYPES
PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
             Sequence 77, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION
GENERAL INFORMATION:
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from
TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
TITLE OF INVENTION: applications thereof
FILE REFERENCE: B4809A JA2
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 137; DB 10;
Pred. No. 2.9e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY,AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REGISTRATION NUMBER: 29,775
REFERRENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF
AND THEIR USE AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 14, Application US/0955138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Hepatitis C virus 7-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                         NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Juery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
New sequences of hepatitis C virus generates for diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
APPLICANT:
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
                       Gaps
                     ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                              TITLE OF INVENTION: genotypes for diagnosis, prophylax NUMBER OF SEQUENCES: 270
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                   5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.5%; Score 137; DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 4.7e-11;
1; Mismatches 2
                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKTKKKAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                              US-09-899-046-152; Sequence 152. Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 152, Application US/09878281.
Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 166 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 166 amino acids
amino acid
               25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
               Matches
                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                        Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                    Score 137; DB 10;
Pred. No. 3.9e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 137; DB 10;
Pred. No. 3.9e-11;
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                          ATORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                              TOPOLOGY: 11near;
MOLECULE TYPE: peptide;
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                  89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 89.37
"---hes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                          3
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 169;
        Length 169
                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
  Score 137; DB 11;
Pred. No. 4.8e-11;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 137; DB 11;
Pred. No. 4.8e-11;
1; Mismatches 2;
                                                                                      1 PKPORKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                 S PKPQKKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                  Sequence 42, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 44, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
    89.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89.5%;
ilarity 89.3%;
Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                               Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : 169 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            169 amino acids
amino acid
Query Match
Best Local Similarity 89.3%
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             , MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE:
US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                              RESULT 14
US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                        ô
                                                                                                                              윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 11; Length 169;
                        2; Indels
                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: TEM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
Pred. No. 4.7e-11;
1; Mismatches 2
                                                              1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11arity 89.3%; Score 137; 12arity 89.3%; Pred. No. 4. Conservative 1; Mismatch
                                                                                      5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRKAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 44, Application US/09899046 Ublication No. US20030008274A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : 169 amino acids
amino acid
                      25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              inear
  Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                          US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-899-046-44
                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
Query Match
Best Local Similarity 89.3%; Pred. No. 4.8e-11;
Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

Q Q

Search completed: August 7, 2003, 12:01:15 Job time: 15.3636 secs

Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Database

Result

Ş

450/80

```
Hepatitis C virus
NC mosaic protein
Hepatitis C virus
Blood transmiscibl
HCV-SI full-length
Hepatitis C virus
Non-A, Non-B Hepat
Non-A, Non-B Hepat
Hepatitis C virus
Recombinant protei
CNI4 fragment of H
HCV capsid peptide
HCV core-envelope
HCV core-envelope
HCV core-envelope
HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Non-A, Non-B Hepa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NC mosaic protein amino acid fragment J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                AAW66083
AAY26952
                                                                                                                                                                                                                                                                                                        AAR54065
AAW37380
                                                                                                                                                                                                                                                                                                                                                           AAY94410
AAY94409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY06682 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US17385.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fields HA, Khudyakov YE;
                                                                           470
28
3010
3010
                                                                                                                                                   WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
 25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  409910506-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-MAR-1999.
AAY06682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY06682
  NC mosaic protein
Hepatitis C virus
Encoded by Hepatit
Encoded by Hepatit
Encoded by Hepatit
NC mosaic protein
HCV type 3 capsid
Hepatitis C virus
Hepatitis C virus
                                                                                 Search time 38.5455 Seconds (without alignments)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                            (SIDSI)/gcgdata/geneseqy/eneseqp-embl/AA1980.DAT:*
(SIDSI)/gcgdata/geneseqy/geneseqp-embl/AA1981.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1981.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1982.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1982.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1986.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1986.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1986.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1980.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1980.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1980.DAT:*
(SIDSI)/gcgdata/geneseqy-embl/AA1991.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                  .al number of hits satisfying chosen parameters:
                                                                                                                                                 149
1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                           1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                    August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                            protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR34473
AAR34474
AAR34475
AAY06673
AAB71258
AAR92968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY06682
AAR94462
                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                     US-09-491-146A-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20
114
120
171
171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ВВ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3023
470
470
470
189
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0
87.2
87.2
87.2
84.6
84.6
84.6
```

셤

ò

```
Gaps
                                                                                                                                                                 Gaps
(HCV) genomic RNA. The protein can be easily detected by antibodies in an assay for the detection of HCV. The DNA and the protein are useful for classifying the subtype of HCV. At least a part of the protein may be used as a vaccine against HCV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B 9405 nuclectides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-A. Primer/probes derived from th sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                             ó
                                                                                                                             Length 3023;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA and cDNA of hepatitis C virus – useful as probes for
diagnosing HCV infection
                                                                                                                         100.0%; Score 149; DB 17;
100.0%; Pred. No. 7e-12;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87.2%; Score 130; DB 14;
82.1%; Pred. No. 4.8e-10;
ive 4; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Encoded by Hepatitis C virus clone JK3-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                    5 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 4; Page 26-28; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                              Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                          AAR34473 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR34474 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 82.1.,
Annes 23; Conservative
                                                                                                      Ouery Match
Ouery Local Similarity 100...
Best Local Similarity 100...
                                                                                                                                                                                                                                                                                                                                                                                                30-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       See AAQ40425-Q40439.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1993-130638/16.
                                                                                       3023 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAQ40431.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-MAR-1993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR34474;
                                                                                                                                                                                                                                                                                                                                                              AAR34473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seguence
                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 4
AAR34474
ID AAR3
XX
AC AAR3
XX
DT 30-J
                                                                                                                                                                                                                                                                                        RESULT 3
                                                                                                                                                                                                                                                                                                         AAR34473
QQ
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                             The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial an mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis. The artificial gene and the resulting mosaic protein improve the sensitivity, specirum of immunoreactivity, and antigen specificity of enzyme spectrum of immunosasays. This provides improved detection of hepatitis C virus. Sequences AAV06673-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence is a polypeptide comprising a 3023 amino acid sequence which is encoded by a cDNA (AAT13279) to a hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus genomic RNA, DNA and related proteins - useful for detection, diagnosis and identification of hepatitis C virus sub-type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1505..1520
/note= "this part of the sequence is missing from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "this part of the sequence is missing from the specification"
                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                              Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis C virus; antibody; detection; diagnosis; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                          Score 149; DB 20;
Pred. No. 5.2e-14;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the specification" 2433..2448
                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 4; Page 16-23; 25pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR94462 standard; Protein; 3023 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                            100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus polypeptide.
               Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94JP-0223933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94JP-0223933
                                                                                                                                                                                                                                                                                                                                               Tuery Match states states 100.0 these 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1996-182301/19.
N-PSDB; AAT13279.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             subtype.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SAYA/) SAYAMA K.
                                                                                                                                                                                                                                                                                                                          28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JP08056672-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-SEP-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       05-MAR-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                classify;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR94462;
                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pept1de
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR94462
ID AAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      é
```

ö

Gaps

ö

Indels

28

Length 470;

```
CDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-C. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-Q40439.
                                                                                                          Score 130; DB 14;
Pred. No. 4.8e-10;
4; Mismatches 1;
                                                                                                                                                             84.6%; Score 126;
                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
         Claim 4; Page 30-32; 44pp; Japanese.
                                                                                                                                                                                                                       Ā
                                                                                                                                                                                                                      AAY06673 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 5; Fig 9; 66pp; English
                                                                                                          Query Match
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                               97US-0921887
                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Khudyakov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-204671/17;
                                                                                        470 AA;
                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28 AA;
                                                                                                                                                                                                                                                                                                                                                   WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                           21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                               25-AUG-1997;
                                                                                                                                                                                                                                                             17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                        04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fields HA,
                                                                                        Sequence
                                                                                                                                                                                                                                        AAY06673;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                 g
                                                                                                                                                                                                                                          * 8 × 5 5 5 5 5 5 5 × 8
                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                            cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-B. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                           Length 470;
                             NANBHV; liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    non-B hepatitis virus; NANBHV; liver disease;
                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA and cDNA of hepatitis C virus – useful as probes for diagnosing HCV infection
                                                                                                                                                                                                      DNA and cDNA of hepatitis C virus - useful as probes for diagnosing \ensuremath{\mathsf{HCV}} infection
                                                                                                                                                                                                                                                                                                                                          Score 130; DB 14;
Pred. No. 4.8e-10;
4; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polymerase chain reaction; diagnostic method
                             HCV; non-A, non-B hepatitis virus; NANBHV; 1
polymerase chain reaction; diagnostic method
                                                                                                                                                                                                                                                                                                                                                                                Encoded by Hepatitis C virus clone JK3-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Encoded by Hepatitis C virus clone JK3-C
                                                                                                                                                                                                                                        Claim 4; Page 28-30; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Æ
                                                                                                                                                           (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR34475 standard; Protein; 470
                                                                                                                                                                                                                                                                                                                                           87.2%;
82.1%;
                                                                                                                      91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91JP-0153736
                                                                                                                                       91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 82.1'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1993-130638/16
                                                                                                                                                                               WPI; 1993-130638/16
                                                                                                                                                                                                                                                                                                    See AAQ40425-Q40439
                                                                                                                                                                                                                                                                                                                       470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAQ40433,
                                                                                                                                                                                         N-PSDB; AA040432
                                                                             JP05068562-A.
                                                                                                                     30-MAY-1991;
                                                                                                                                      30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-MAY-1991;
                                                                                                23-MAR-1993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR34475;
                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR34475
ID AAR3
XX
                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The nortigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation. The method for protein and the artificial mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against mosaic gene and protein is also useful for vaccination against artificial gene and protein is also useful for vaccination against artificial gene and the resulting mosaic protein improve the sensitivity, specially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, specially mosaic protein and artigen specificity of enzyme communoassays. This provides improved detection of hepatitis C virus.
                                                                                                                             Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             antigenic
for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New mosaic protein, comprising a plurality of homologous peptides from different genotypes of a species - useful detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (USSH ) US DEPT HEALTH & HUMAN SERVICES.
mosaic protein amino acid fragment
```

Length 28;

DB 20;

```
AAR92968 standard; Protein; 191 AA.
                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                        Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         191 AA;
                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAT16642
                                                                                                                                                                                                                            WO9605315-A2.
                                                                                                                                                                                                                                                                                                         .5-AUG-1994;
                                                                                                                                                                                                                                                                                .5-AUG-1995;
                                                                                                           02-0CT-1996
                                                                                                                                                                                                                                                      22-FEB-1996.
                                                                                                                                                                           hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis.
                                                                                AAR92968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR92969;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                             RESULT 8
AAR92968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This invention describes a novel attenuated flavivirus live vaccine comprising a flavivirus mutant that has a deletion of at least 4 consecutive amino acids from the capsid protein, provided that the consecutive amino acids from the capsid protein, provided that the vaccine of the invention has virucide, antinflammatory and hepatotropic vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a encephalitis, dengue, classical swine fever, bovine viral diarrhoea and hepatitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does of effect immune responses to important proteins; and (iii) can not generate a non-natural virus by recombination. The mutant viruses eliminate the need to produce large amounts of infectious/virulent viruses, and can be produced with less expense. The protective response to flavivirus lasts significantly longer than that to killed vaccines. This sequence represents a fragment of the capsid protein from Hepatitis curver.
                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids from the capsid protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                   Gaps
                 ö
                                                                                                                                                                                                                                    Capsid protein; attenuated vaccine; virucide; antinflammatory; hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 189;
                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
Best Local Similarity 78.6%; Pred. No. 9.1e-11;
Matches 22; Conservative 5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84.6%; Score 126; DB 23;
ilarity 78.6%; Pred. No. 6.8e-10;
Conservative 5; Mismatches 1;
                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPORKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                  Ą.
                                                                                                                                                                                                             HCV type 3 capsid protein fragment.
                                                                                                                             AAB71258 standard; protein; 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 2; 30pp; German.
                                                                                                                                                                                                                                                                                                                                                                    11-FEB-2002; 2002WO-AT00046
                                                                                                                                                                                                                                                                                                                                                                                              21-FEB-2001; 2001AT-0000272.
                                                                                                                                                                                18-NOV-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Heinz FX, Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                      (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                                                                                                                                        depatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             189 AA;
                                                                                                                                                                                                                                                                                                                WO200266621-A1.
                                                                                                                                                                                                                                                                                                                                           39-AUG-2002.
                                                                                                                                                        AAB71258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                  RESULT 7
                                                               셤
                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          388888888888888
```

```
HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunishing against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, Lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; El; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection % \left( 1\right) =0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 126; DB 17;
Pred. No. 6.8e-10;
Hepatitis C virus isolate HK10 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus isolate S52 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 6.86
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR92969 standard; Protein; 191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 4; Page 207; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Purcell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                              94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02-OCT-1996 (first entry)
```

ò g

Gaps

Fri Aug

388888888**%**8

```
o acid in this position is designated X in specification, but codon usage shows that the possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid in this position is designated X in
the specification, but codon usage shows that the
only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection.
                                                                 AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounchear cells. The antibodies may be used in the prevention of HCV infection.
      to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317.
                                                                                                                                                                                                                                                    Length 191;
                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                               Score 126; DB 17;
Pred. No. 6.8e-10;
5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /label- Gly
/note- "amino acid in this po
the specification, b
only possible amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "amino acid in this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 144..149
/label- Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Met, Thr, Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         G_{1}y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label- Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ile, Val
                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                               Claim 4; Page 209-210; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label- Val, Ala, Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                             AAR96547 standard; peptide; 319 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label Cys, Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label - Met, Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'label- Ser, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Leu
                                                                                                                                                                                                                                                 84.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label= Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                              Local Similarity 78.6 ies 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 177
                                                                                                                                                                                                                 191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-MAR-1997
                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR96547;
                                                                                                                                                                                                                                               Query Match
      nseq
                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                            AAR96547
      a
                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                               AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                      DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA and amino acid sequence of HCV envelope 1 and core proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 126; DB 17; Length 191;
Pred. No. 6.8e-10;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus isolate DK12 core protein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                     Claim 4; Page 208; 340pp; English
                                                                                                                                                                       Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR92971 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.6%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Miller RH, Purcell
                                                            95WO-US10398
                                                                                            94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                       WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                        N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAT16645
WO9605315-A2
                                                                                            15-AUG-1994;
                                                            15-AUG-1995;
                              22-FEB-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-FEB-1996,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR92971;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; El;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bukh J,
                                                                                                                                                                         Bukh J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
AAR92971
ID AAR92
XX
AC AAR92
```

g ò

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein a lasto useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAV06673-681 sepresent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                   New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus type 11a isolate FR19 amino acids 1-74.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.9%; Score 125; DB 20;
82.1%; Pred. No. 1.3e-10;
11ve 3; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                            (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR96577 standard; peptide; 74 AA.
                                                                                                                                                                                                                                       Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95EP-0870076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95WO-EP04155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 82.1
Matches 23; Conservative
                                                                    Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Maertens G, Stuyver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1996-251460/25.
N-PSDB; AAT27988.
                                                                                                           WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUN-1995;
21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9613590-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14-MAR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-MAY-1996.
                                                                  Fields HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR96577;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR9657
        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hepatitis C virus subtypes 1 c. 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes 1a-c, 2a-d, 3a-f, 3g, 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5' genome. This sequence represents amino acids 1-317 from the HCV types 7c and 8a isolates VN4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The new HCV types were isolated from patients with chronic HCV from the benelux countries. France, Cameroon and Vietnam, because of their aberiant reactivities. The RNA was extracted, cDNA synthesised and PCR amplified, cloned and genotyped. The YUR, Core/El and NSSB regions were sequenced either directly or partially and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences were used to generate the peptides AAR96424-R96524. The aequences can also be used to synthesise probes and primers for the detection of HCV in a sample. The polypeptides can be used codetect anti-HCV antibodies, for HCV typing or to prevent HCV infections.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequences AAR96526-R96578 represent novel sequences isolated from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        84.6%; Score 126; DB 17; Length 319; 78.6%; Pred. No. 1.2e-09; 1ve 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        restriction endonuclease assisted ligation; vaccination
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NC mosaic protein amino acid fragment B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY06674 standard; Protein; 28 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 25; Fig 3; 150pp; English.
                                                                                 95WO-EP04155
                                                                                                                         95EP-0870076
94EP-0870166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 78.6 les 22; Conservative
                                                                                                                                                                                        (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                                                        WPI; 1996-251460/25.
N-PSDB; AAT27958.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  319 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
WO9613590-A2
                                                                                                                         28-JUN-1995;
21-OCT-1994;
                                                                                 23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-AUG-1997;
                                         09-MAY-1996
                                                                                                                                                                                                                                Maertens G,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY06674;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hesULT 12
AAY06674
```

ò g ö

Gaps

against HCV in serum, The antibodies may

ö

Gaps

ö

Indels

Length 191;

```
HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                  Score 125; DB 17;
Pred. No. 9.5e-10;
3; Mismatches 2;
proteins may also be used to detect antibodies a saliva, lymphocytes or other mononuclear cells. be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus isolate S2 core protein.
                                                                                                                                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US DEPT HEALTH & HUMAN SERVICES.
US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: August 7, 2003, 11:14:10
Job time : 39.6364 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR92970 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 4; Page 209; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Purcell RH;
                                                                                                                ch 83.9%;
1 Similarity 82.1%;
23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 83.9%;
75.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 83.9
Best Local Similarity 75.0
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1996-139709/14
                                                                                                                Query Match
Best Local Similarity
Matches 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA;
                                                                              Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAT16644
                                                                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                         02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis.
                                                                                                                                                                                                                                                                                                                                                                    AAR92970;
                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (USSH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bukh J,
                                                                                                                                                                                                                                                                                        RESULT 15
AAR92970
                                                                                                                                                                                                                                                                                                                                                                    ្តី
ស្ត្រី
ស្ត្រី
ស្ត្រី
                                                                                                                                                                                             ð
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                     Benelux countries, France, Cameron and Vietnam, because of their aberrant reactivities. The RNA was extracted, cDNA synthesised and PCR amplified, cloned and genotyped. The 5 UK. Core-fin and NSSB regions were sequenced either directly or partially and used to classify the new viruses into (sub)types based on comparison with known sequences. The sequences were used to generate the peptides ARA96424-R96524. The sequences can also be used to synthesise probes and primars for the detection of HCV in a sample. The polypeptides can be used to detect anti-HCV antibodles, for HCV typing or to prevent HCV infections.
                                            The sequences AAR96526-R96578 represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes ld-f, 2e-i, 2k, 2l, 3g, 4k-m, 7a-c or types 9, 10 or 11. The sequences corresp. to the 5' untranslated region (UR), the Core/El, NS4 or NS5B regions of the genome. This sequence represents amino acids 1-74 from the HCV type lla isolate FR19.
                                                                                                                                                                                         The new HCV types were isolated from patients with chronic HCV from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    envelope 1; core protein; MCV genotyping; antibody; vaccine,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 125; DB 17;
Pred. No. 3.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus isolate S45 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 4; Page 186-187; 340pp; English.
                Claim 25; Fig 3; 150pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92943 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                         83.9%;
82.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94US-0290665.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                       74 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT16617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR92943;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
AAR92943
ID AAR92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
```

ž

à g

\*&\$999999999999999

```
ö
AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, sallva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                          Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                  Score 125; DB 17;
Pred. No. 9.5e-10;
6; Mismatches 1;
```

us-09-491-146a-32.rag

```
August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
GenCore version 5.1.6
Copyright (c) 1993 · 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                al number of hits satisfying chosen parameters:
                                                                                                                                                                                                             149
1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                283308 seqs, 96168682 residues

    protein search, using sw model

                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                           US-09-491-146A-32
                                                                                                                                                                                         Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                       Scoring table:
                                                                          OM protein
                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                Run on:
```

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Minimum DB seq length: 0 Maximum DB seq length: 200000000 Database :

PIR\_76:\*
1: pir1:\*
2: pir2:\*
3: pir3:\* pir4:\* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES Res

		:	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein	tein
	4	C10n	polypro	polypro	polypro	polypro	polyprotein	polypro	polyprotein	polypro	polypro	polyprotein	polyprotein	polyprotein	polyprotein	polypro	polypro	polypro	polyprotein	polypro	polyprotein	polypro	polyprotein								
	1	nescription	депоше																				genome			genome			enome		
2																															
			175	090	159	58	091	191	88	151	149	84	. 951	:73	136	191	153	155	157	148	171	141	170	69	. 89	142	344	350	154	145	4
J	£	3 :	S198	PC20	S413	S413	S41360	PC20	8412	S413	S413	3015	S413	A455	S21336	S413	S413	S413	8413	5413	S413	S413	S413	S413	S413	S413	41	47	Н	8413	S413
	9	9 :	7	7	~	N	~	~	7	7	7	~	~	~	~	~	~	~	~	7	ď	7	~	N	~	~	ď	~	~	~	7
	10001		782	411	114	114	124	411	492	115	115	640	108	3010	88	123	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115
æ	Query														81.9																
	0,000	2000	130	129	126	126	126	126	126	125	125	125	123	123	122	122	121	121	121	121	121	121	121	121	121	121	121	121	121	121	121
	ssult	2 :	Н	7	m	4	'n	9	7	œ	σ	10	11	12	13	14	15	16	17	18	19	50	21	7.	23	54	25	56	27	28	53

genome polyprotein genome polyprotein genome polyprotein genome polyprotein genome polyprotein polyprotein polyprotein polyprotein genome polyprotein genome polyprotein genome polyprotein genome polyprotein genome polyprotein genome polyprotein pypothetical prote genome polyprotein	ALIGNMENTS	C virus (isolate JK3) (fragment) ope protein 1; nonstructural protein 2; NS1/E2 prot	30-Jun-1992 #text_change 17-Nov-2000	.; Kobayashi, K.; Murakami, S. September 1991 putative structural regions of Hepatitis C Vir	482: PTDN.CBA43789 1. PTD.459483	orotein redope protein; gl ed cMAT1> predicted cMAT2> sdicted cMAT3> fragment) #status	re 130; DB 2; Length 782; d. No. 4.1e-10; Mismatches 1; Indels 0; Gaps 0;
S41343 S24134 S12471 S12707 S12707 PC1284 J01926 J10711 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032 S18032	ALIG	virus ( e prote	ision 3	U.; Kol y, Septi of puta	975 CTN	genome pol protein; e atus predi in 1 #stat #status p	Score 130; Pred. No. 4; Mismatc
новововововов		၁၀	rev	hi, rar is		"JX us re #st ote	00 d0
1115 1118 3618 3613 5520 7882 7882 7874 874 874		.n - hepatitis C virus (iso protein; envelope protein Lis C virus	#sequence_revision	S.; Masashi, U Data Library, Ce analysis of	1c RNA	isolate isolate tein; co protein elope pr /E2 prot	87.2%; larity 82.1%; Conservative
addadddddddddddd		ote	# #	SI SI SI	E ON ME	ce: cre pro re re NS1	rit
		in Pro	992 875	eko EM equ	975 9ei 2, <i< td=""><td>courcepant of the cource cource cource cource cource tr: fr: fr: fr: fr: fr: fr: fr: fr: fr: f</td><td>11a) Coi</td></i<>	courcepant of the cource cource cource cource cource tr: fr: fr: fr: fr: fr: fr: fr: fr: fr: f	11a) Coi
		ESULT 1 19875 enome polyprotein - hepatit contains: core protein; en ;Species: hepatitis C virus varietu : solate IK3	C; Date: 30-Jun-1992 C; Accession: S19875	R; Honda, M.; Kaneko, S.; Masashi, U submitted to the EMBL Data Library, A; Description: Sequence analysis of A; Reference number: S18029	A; Accession: S19875 A; Molecule type: genomic R A; Residues: 1-782 <hon> A; Cross-references: FMMI.: X</hon>	A. Experimental source: isolate JK3 C; Superfamily: hepatitis C virus genome polyy C; Keywords: capsid protein; core protein; en: F;1-191/Product: core protein #status predict F:192-383/Product: envelope protein 1 #status F;384-733/Product: NS1/E2 protein #status pre F;734-782/Product: nonstructural protein 2 (1	Query Match Best Local Similarity Matches 23; Conser
		RESULT 1 S19875 genome polyi N;Contains: C;Species: B	C; Date C; Acce	R; Hond submit A; Desc A; Refe	A; Resident A; Resident A; Resident A;	A) EXPO C) Suppo C) Suppo F) 1-19 F) 384-17	Query M Best Lo Matches

Procession Parallels C virus
N; Contains: envelope protein E1; nonstructural protein E2/NS1
C; Specials: hepatitis C virus
C; Specials: hepatitis C virus
C; Specials: hepatitis C virus
C; Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 17-Nov-2000
C; Accession: Pc2060
R; Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Blochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A; Tille: Identification of the third major genotype of hepatitis C virus in France
A; Reference number: Pc2060; MuID:94197744; PMID:8147893
A; Residues: 1-411 < LIJ>
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capabid protein; envelope protein; glycoprotein; nonstructural pro
F; 192-383/Product: envelope protein E1 #status predicted <NPE>
F; 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) #status predicte 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28 δ 음

```
541288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ground Purprotects

William Contains: Core protein

CiSpecies: hepatitis C virus

Aivariety; genotype 3, N1

CiSpecies: hepatitis C virus

Aivariety; genotype 3, N1

CiDate: 19-May 1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000

CiAccession: 541358

Titled to the EMBL Data Library, January 1994

Ascription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A; Reference number: S41318

A; Accession: 841358

A; Accession: 841358

A; Accession: 841358

A; Molecule type: genomic RNA

A; Residues: 1-114 < vANN

A; Residues: 1-114 < vANN

A; Residues: 1-114 < vANN

A; Cross-references: EMBL: 229461; NID: 9443884; PIDN: CAA82599.1; PID: 9443885

A; Experimental source: genotype 3, NI

C; Superfamily: hepatitis C virus genome polyprotein

C; Keywords: capsid protein; core protein; polyprotein

F; 1-114/Product: core protein #status predicted < vAT>
                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                       C; Species: hepatitis C virus
A; Variety: genotype 3, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S4,1359
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
'Sscription: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
iference number: S41341
                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: EMBL:229462; NID:9443886; PIDN:CAA82600.1; PID:9443887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 3, N2) (fragment)
N.Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein - hepatitis C virus (genotype 3, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S41360
genome polyprotein – hepatitis C virus (genotype 3, N3) (fragment)
       Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 114;
                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A:Experimental source: genotype 3, N2
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-114/Product: core protein *status predicted <MAT>
Score 129; DB 2;
Pred. No. 3.1e-10;
5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match

84.6%; Score 126; DB 2;
Best Local Similarity 78.6%; Pred. No. 2.3e-10;
Matches 22; Conservative 5; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84.6%; Score 126; DB 2; 78.6%; Pred. No. 2.3e-10; iive 5; Mismatches 1
                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              86.6%;
78.6%;
                                                    22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A: Residues: 1-114 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 22; Conserv
                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Accession: $41359
                          Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                 ò
                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
genome polyprotein N2 - hepatitis C virus
N;Contains: envelope protein E1; nonstructural protein E2/NS1
N;Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Species: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
C;Accession: PC2061
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
A;Itle: ldentification of the third major genotype of hepatitis C virus in France.
A;Reference number: PC2060; MUID:94197744; PMID:8147893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     prot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1-492 <SEE>
A; Cross-references: EMBL:X76918
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; envelope protein; nonstructural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A) Cross-references: GB:L12355, NID:g410169, PIDN:AAA.20155.1, PID:g410170
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; glycoprotein; nonstructural
F; 192-383/Product: envelope protein El #status predicted <SPE>
F; 384-411/Product: nonstructural protein E2/NS1 #status predicted <NPE>
F; 196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predi
N; Contains: core protein
C; Species: hepatitis C virus
C; Species: hepatitis C virus
A; Varlety: genotype 3, N3
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S4136
B; van Doorn, L.J.; Rleter, G.E.M.; Brouwer, J.T.
A; Date: 19-May-1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Accession: S4136
A; Accession:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NS1 protein
C;Species: hepatitis C virus
C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A)Cross-references: EMEL:229463; NID:9443888; PIDN:CAA82601.1; PID:9443889
A)Experimental source: genotype 3, N3
C;Superfamily: hepartitis C virus genome polyprotein
C;Keywords: capaid protein; core protein; polyprotein
F;1-124/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 126; DB 2;
Pred. No. 2.5e-10;
5; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 126; DB 2
Pred. No. 8e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Data Library, December 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.68;
78.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Accession: S41288
R;Seelig, R.
submitted to the EMBL Data L
A;Reference number: S41288
A;Accession: S41288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches , 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 1-411 <LIJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

2

```
÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (genotype 1, N7) (fragment)
N;Contains: core protein
Species: hepatitis C virus
ariety: genotype 1, N7
Late: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                  N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N9
Pate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Molecule type: genomic RNA
A;Residues: 1-115 <VAN>
A;Cross-references: EMBL:229452; NID:g443866; PIDN:CAA82590.1; PID:g443867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C;Accession: S41349
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
F;1-191/Product: core protein #status predicted <COR>
F:192-372/Product: envelope protein #status predicted <ENV>
F:373-492/Product: NSI protein (fragment) #status predicted <NSI>
                                                                                                                                                                                                                                                                                                                                                 polyprotein - hepatitis C virus (genotype 1, N9) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 115;
                                                                                            Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 115,
                                                                                          Score 126; DB 2; Length 49
Pred. No. 9.5e-10;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 125, DB 2, Length 11 Pred. No. 3.2e-10, 3. Mismatches 2, Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A; Description: Analysis of hepatitis C virus genotypes I to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 125; DB 2; L
Pred. No. 3.2e-10;
3; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A:Experimental source: genotype 1, N9
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein *status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A:Experimental source: genotype 1, N7
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKROTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83.9%;
llarity 82.1%;
Conservative
                                                                                       Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity 82.1%;
23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Cross-references: EMBL: 229454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
es 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                              .ccession: S41351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                      ö
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
```

```
genome polyprotein - hepatitis C virus (strain JT) N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nons protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cispecies: hepatitis C virus
Cispecies: hepatitis C virus
Cispecies: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
Cispecession: A45573
R;Tannaka, T; Kato, N; Nakagawa, M; Ootsuyama, Y; Cho, M.J.; Nakazawa, T.; Hijl Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carr
A;Reference number: A45573; MUID:92295714; PMID:1318627
                                                 nonstructura
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nicontains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N5
A; Variety: genotype 2, N5
C; Accession: S41356
B; Var Doorn, L.J.; Rieter, G.E.M.; Brouwer, J.T.
R; Var Doorn, L.J.; Rieter, G.E.M.; Brouwer, J.T.
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Accession: S4136
A; Accession: S4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
       tis C virus (strain U.K.) (fragment)
envelope protein El; envelope protein E2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.9%; Score 125; DB 2; Length 64
82.1%; Pred. No. 1.7e-09;
11ve 3; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A:Experimental source: genotype 2, N5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 123; DB 2;
Pred. No. 5.7e-10;
1; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82.6%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                larity 82.1%;
Conservative
       hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                          N;Contains: core protein C; e
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
nes 23; Conserv
genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ß
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

Gaps

ö

```
C; Species: hepatitis C virus
A; Variety; genotype 2, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41353
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipp.
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
                                                                                                                                                         DB 2; Length 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108,
                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 121; DB 2; L
Pred. No. 1.1e-09;
4; Mismatches 2;
              A; Experimental source: genotype 3, N4
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-123/Product: core protein #status predicted AMI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: EMBL:229456
A;Experimental source: genotype 2, N2
C;Superfamily: hepatitls C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <MAI>
                                                                                                                                                      Score 122; DB 2;
Pred. No. 9e-10;
6; Mismatches
                                                                                                                                                                                                                                                                                          1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7, 2003, 11:21:49
                                                                                                                                                      81.9%;
75.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                          Ouery Match
Best Local Similarity 75.0°
Matches 21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search completed: August
Job time: 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N; Contains:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                    A: Experimental source: HCV-JT
A: Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C: Superfamily: hepatitis C virus genome polyprotein
C: Superfamily: hepatitis C virus predicted <PCP>
F: 2-115/Product: capsid protein C *status predicted <PCP>
F: 116-19/Product: nonstructural protein NS1 *status predicted <NS1>
F: 300-729/Product: nonstructural protein NS2 *status predicted <NS2>
F: 1007-1615/Product: hepativirin *status predicted <NS3>
F: 1230-1237/Region: nucleotide-binding motif B (P-loop)
F: 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             submitted to the EMBL Data Library, April 1992
A;Description: A sentitive serodiagnosis of hepatitis C virus infection with two cloned
A;Reference number: S21336
                                                                                   Residues: 1-3010 <TAN>
Cross-references: GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 3, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) C;Species: hepatitis C virus (C.Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C;Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;MOlecule type: genomic RNA
A;Residues: 1-123 <VAN>
A;Cross-references: EMBL:229464; NID:g443890; PIDN:CAA82602.1; PID:g443891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NID:959492; PIDN:CAA46517.1; PID:959493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F.1316-1319/Region: DEXH motif
F.1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F.1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
'014-3010/Product: nonstructural protein NS5 #status predicted <N85>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (genotype 3, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 123; DB 1; Length 30
Pred. No. 1.4e-08;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A:Description: Analysis of hepatitis C virus genotypes 1 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.9%; Score 122; DB 2; 78.6%; Pred. No. 6.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  y Match
Local Similarity 78.6%;
hes 22; Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A)Cross-references: EMBL:X65548;
C;Superfamily: hepatitis C virus
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    est Local Similarity 78.6 atches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA
A; Residues: 1-88 <SAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Reference number: S41341
                                  Status: preliminary.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Accession: S41361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Juery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
A;Accession:
A;Status: pre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

ö

Gaps

	Ltd.
5.1.6	Compugen
version	- 2003
GenCore	(c) 1993
	Copyright

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments)
268.226 Million cell updates/sec Run on:

US-09-491-146A-32 149 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 :al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description		P2/959 hepatitis C O01404 hepatitis C			욛.	у депоше	P2000z n genome po P27958 h genome po	h qenome	h genome	h genome						-		-	000725 drosophila	Q9xfs9 arabidopsis	Q9ugu0 homo sapien	_	P47915 mus musculu	_	Q9p0k1 homo sapien	-	Q08865 volvox cart	000515 homo sapien		_	Q9kh11 thermus aqu
SOMMANIES	POLG_HCVJT	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVH	POLG_HCVJ6	POLC_HCVJ8	POLG_HCVTW	POLG_HCV1	PIII_MOUSE	POLG_BVDVN	H1_LYCPN	FURI_HUMAN	DYHC_DICDI	H12_RAT	RL29_YEAST	SGS4_DROME	DXR_ARATH	TF20_HUMAN	TF20_MOUSE	RL29_MOUSE	VATA_METTH	AD22_HUMAN	POLS_WEEV	H12_VOLCA	LAD1_HUMAN	POLG_BVDVS	RL5_RABIT	DTD_THEAQ
DB		٦,	7	<b>~</b>	٦,			٦	-	-	٠,	<b>-</b>	٦,	٦.	٦.	<b>.</b>	, I	<b>,</b> —1	٦,		н,	٠,		<b>.</b>	-	<del></del> 1	<del>, -</del> 1	<del>, -</del> 1			~
% Query Match Length	3010	520	520	737	737	3010	3011	3033	3033	3010	3011	332	3988	202	794	4725	218	28	297	477	1960	1983	159	584	906	1236	240	517	3898	101	152
	: 2:	81.2	ä	ä,	∹,	i.	81.2	ä	ä	٠	75.8	4	4 (		m.	N	•		31.5		31.5	Ξ,	ή.	٠	;	;	30.9	·	·	•	Ö
Score	123	121	121	121	121	121	121	121	121	114	113	25	51.5	• •	20	٠,	47.5	47	47	47	47		46.5	٠.	ė.		46	4		42	45
Result No.		n w	4	,	o r	~ a	9 6	10	11	12	13	<b>7</b> 7	12	91	17	87.	19	50	21	7.7	23	77	25	26	27	28	53	30	31	32	33

H1_SALTR RL5_MOUSE RL5_RAT RL5_RAT RT6_RAT RT7_MOUSE FURI_MOUSE FWF9_YENAT PPAN_ARATH COCH_HUMAN SYKC_YEAST FUKA_C_CADBI FUKI_BOVIN	P02254 salmo trutt	P4/962 mus musculu	P09895 rattus norv	P29673 drosophila	P23188 mus musculu	P21573 xenopus lae	Q04958 saccharomyc	O9asu7 arabidopsis	043405 homo sapten	P15180 saccharomyc	045882 clostridium	Q28193 bos taurus
	H1_SALTR	KL2_MOUSE	RL5_RAT	APTE_DROME	FURI_MOUSE	YB1_XENLA	YMF9_YEAST	PPAN_ARATH	COCH_HUMAN	SYKC_YEAST	CGAA_CLOBI	FURI_BOVIN
	4 4	D (	296	469	793	303	1679	345	550	290	613	197
194 296 296 296 303 303 1679 345 590 590 797	516											
		4.0	~	7	۲.	6.	6.	'n.	ς.		٦.	ń
30.2 194 30.2 296 30.2 296 30.2 296 30.2 793 29.9 1679 29.5 345 29.5 550 29.5 590		7.00	30.2	30.2	30.2	29.9	29.9	29.5	29.5	29.5	29.5	29.5
		7.05	45 30.2	45 30.2					44 29.5	44 29.5	44 29.5	44 29.5

## ALIGNMENTS

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E)
MAJOR ENVELOPE PROTEIN NSJ. F2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NSJ (POTENTIAL).
NON-STRUCTURAL PROTEIN NSJ (POTENTIAL).
NON-STRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 123; DB 1; Length 3010;
Pred. No. 1e-09;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.

N-LINKED (GLCNAC.)

N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                    InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR0071095; RNA_DOL_DS_FS.
InterPro; IPR007094; RNA_DOL_DS_FS.
InterPro; IPR007094; RNA_DOL_DS_FS.
InterPro; IPR007094; RNA_DOL_DS_FS.
InterPro; IPR007094; RNA_DOL_DS_FS.
InterPro; IPR0071094; HCV_Caps1d; I.
Ffam; PF01539; HCV_NS1; I.
Ffam; PF01001; HCV_NS2; I.
Ffam; PF01001; HCV_NS3; I.
Ffam; PF01001; HCV_NS4a; I.
Ffam; PF01001; HCV_NS3; I.
Ffam; PF01001; HCV_NS3; I.
Ffam; PF01015; Hellcase_C; I.
Ffam; PF0057; Hellcase_C; I.
Ffam; PF0057; Hellcase_C; I.
Ffam; PF00698; Viral_RAPP; I.
Fr006m; PD186662; HCV_NS1; I.
                           HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82.6%;
milarity 78.6%;
Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        326573
                                                   IPR002518; IPR004109; I
                                      IPR002531;
                                                                                            IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         623
645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                 3D-structure.
                                                   InterPro;
InterPro;
InterPro;
                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropan Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                       MEDLINE=92230222; PubMed=1314459;
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Tomology to reported isolates: comparative study of four distinct agencypes.";
Tenotypes.";
Tenotypes.";
Tenotypes.";
Tenotypes.";
Tenotypes.";
The SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPPOPIEIN ENGLY.
C. I. SUBUNT: THE CONFORTEIN ENGLY OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C.
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
                                                                                                                      protein) (P22);
                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22)
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1)) (Fragment).
Hepatitis C virus (isolate HC-J2) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR00519; HCV_core; 1.
IP fam; PF01349; HCV_core; 1.
IP fam; PF01349; HCV_core; 1.
IP fam; PF01359; HCV_core; 1.
IP F00m; PD186062; HCV_NS1; 1.
IP F00pm; PD186062; HCV_NS1; 1.
IP F019Protein; Glycoprotein; Coat protein; Envelope protein; Iransmembrane; Nonstructural protein.
INII_MET 1 INII_MET 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTAL.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                  513 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D10074; BAA00968.1; -.
                                STANDARD;
                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                          NCBI_TaxID-11111;
                                POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
                  POLG_HCVJ2
RESULT
```

ö

Gaps

; 0

```
POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                                                     Hepacivirus
     NON_TER
SEQUENCE
                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                        Best Loca
Matches
                                                                                                                                                                         POLG_HCVHK
                                                                                                                                                          RESULT
     S F
                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR ANINOPERTIDAGE.
CAPSID PROTEIN C FOTENTAL).
MAJOR ENVELOPE PROTEIN ( FOTENTIAL).
MAJOR ENVELOPE PROTEIN E ( FOTENTIAL).
NONSTRUCTURAL PROTEIN ES! ( FOTENTIAL).
N'LINKED (GLCNAC. . ) ( FOTENTIAL).
                                                                                                                                                                                                                                                                                                                                   MEDLINE-93019030; PubMed-1383400;
MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2739(1992).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRRA.
                                                      Gaps
                                                                                                                                                                                           01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
                                                                                                                                                                                                                                                                      Hepatítis C virus (isolaté HCV-476) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         pfam; pF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein: Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY
                            Length 513;
                                                    Indels
943F31E3514CDEF3 CRC64;
                         Score 121; DB 1;
Pred. No. 3.2e-10;
                                                                                                                                                                      520 AA
                                                   4; Mismatches
                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PP011543; HCV_cappsid; 1.
Pfam; PP011542; HCV_capsid; 1.
Pfam; PP011542; HCV_core; 1.
 55704 MW;
                         81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10688; BAA01530.1; -.
                    Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                    STANDARD;
 513 AA;
                                                                                                                                                                                                                                                                                                               NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1192
3384
1196
1209
2334
424
431
4431
                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                    POLG_HCVH4
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                        RESULT 3
POLG_HCVH4
ID POLG_H
DT 01-JUL
DT 01-JUL
DT 16-OCT
DE GENOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
S
                                                                                                 d
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                            01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GF099 (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN AND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                           Length 520
                                                                                                                        Indels
  520
56499 MW; AA135246CF20D525 CRC64;
                                             81.2%; Score 121; DB 1; Le
78.6%; Pred. No. 3.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BY SIMILARITY.

N-LINKED (GLCNAC.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                         520 AA.
                                                                                                                                                                                            1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR: JO1925; JO1925.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002513; HCV_core.
InterPro: IPR002513; HCV_core.
InterPro: IPR002531; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-93019030; PubMed-1383400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                          Local Similarity 78.6 tes 22; Conservative
520 5
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=31644;
```

```
RESULT
     8
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transmembrane: Nonstructural protein.

INIT_MET 1 1 CELLULAR AMINOPEPTIDASE.

CHAIN 116 191 CAPBID PROTEIN (POTENTAL).

CHAIN 192 383 MAJOR ENVELOPE PROTEIN (POTENTAL).

CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).

CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                  01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
16-OGT-2001 (Rel. 40, Last annotate)
16-OGT-2001 (Rel. 40, Last annotate)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NS2] (Fragment): (HCV)
Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genotypes.;
Virology 188:331-341(1992)

-1- FUNCTION: THE SMALL PROTEINS NS2B, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-1- SUBBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCARSID COVERED BY
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
 N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
                                                                                                           ö
                                                                              Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002512; HCV_core.
InterPro; IPR002513; HCV_env.
InterPro; IPR002531; HCV_capsid; I.
Pfam; PF01542; HCV_capsid; I.
Pfam; PF01542; HCV_care; I.
Pfam; PF01550; HCV_env; I.
ProDom; PF01560; HCV_NS1; I.
ProDom; PF01560; HCV_NS1; I.
ProDom; Pf01560; HCV_NS1; I.
ProDom; Pf01560; HCV_NS1; I.
                                                                 Score 121; DB 1; Length 52. Pred, No. 3.3e-10;
                                                    56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                             737 AA
                                                                                                         4; Mismatches
                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                            32
                                                                                                                                                PKPQRKTKRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                PRT;
                                                                               81.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10075; BAA00969.1; -.
                                                                                            78.68;
                                                                                                         Conservative
                                                                                                                                                                                                                             STANDARD;
424
431
449
520
520 AA;
                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11112;
                                                                                Local Sim.
                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                            POLG_HCVJ5
P27960;
                                        NON_TER
SEQUENCE
 CARBOHYD
             CARBOHYD
                          CARBOHYD
                                                                                Query Match
                                                                                                                                                                                                              POLG_HCVJ5
ID POLG_HC
AC P27960;
DT 01-AUG-
                                                                                                         Matches
                                                                                                                                                                                                 RESULT
 SFFF
                                                                                                                                   ò
                                                                                                                                                         g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0C-2001 (Rel. 40, Last annotate)
16-OCT-2001 (Rel. 40, Last annotate)
Genome polyprotein (Contains: Capation update)
Genome polyprotein (Contains: Capation protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NS2] (Fragment): (HCV).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
-1. FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
-1. FUNCTION: THE SMALL PROTEING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-1. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPHOPIENI ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92230222; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Tril-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct qenotypes.";
                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 121; DB 1; Length 73
Pred. No. 4.7e-10;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3AF699D82AD501B1 CRC64;
                                 N-LINKED (GLCNAC, N-LINKED) (
                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKROTKRNTLRRPKNVKFPAGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_MSI.
Pfam; PF01543; HCV_capsid, 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81207 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJ7
P27961;
                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best
```

```
ö
                                                                                                                                                          CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/F2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
15-5EP-2003 (Rel. 42, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
670-68) (GP70) (NS1); Protein F7; Nonstructural protein RS2 (P21)
62 3.4.22.-); Protease/helicase RS3 (P70) (Hepacivirin)
62 3.4.21.98); Nonstructural protein NS4 (P4); Nonstructural protein NS5 (P50); Nonstructural protein NS5 (P50); Nonstructural protein NS5 (P60); Non
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE-91140698; PubMed-1847440;
Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        · · ) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81.2%; Score 121; DB 1; Length 737; 78.6%; Pred. No. 4.7e-10; 1ve 4; Mismatches 2; Indels
                                                                                               protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA; 81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                  N.LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                    POTENTIAL.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_core; 1.
Pfam; PF01550; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Probom; Glycoprotein; Coat protein Glycoprotein; Coat protein Insamembrane; Nonstructural protein INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 78.6 ies 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                               TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
POLG_HCVBK
ID POLG_H
AC P26663
DT 01-AUG
                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
    Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                      X KAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X MEDLINE-98227846; PubMed-9568891;

X MEDLINE-98227846; PubMed-9568891;

X A Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;

**Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

**T 'Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

**T 'Virus: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'Virus: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'VIRUS: a 2.2-A resolution structure in a hexagonal crystal form.";

**T 'Complex of STATY A ROLE IN THE VIRAL RNA REPLICATION.

**T 'COMPLEX OF STATY IN THE IN and Set or Ala in Pl'.

**T 'COMPLEX OF THY IN PL and Set or Ala in Pl'.

**T 'COMPLEX OF THY IN PL and Set or Ala in Pl'.

**T 'COMPLEX OF THY IN PL and Set or Ala in Pl'.

**T 'COMPLEX OF THY IN PL AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.

**T 'COMPLEX OF THY IN THE AND SET OF Ala in Pl'.
                                                                             X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; PubMed-8861916;
Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a trypain-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROFEIN C AND MRNA.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR002511; HCV_NS2.
InterPro; IPR002109; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR000146; HCV_NS4.
InterPro; IPR001266; HCV_NS5.
InterPro; IPR001469; HCV_NS5.
InterPro; IPR001569; HCV_NS5.
InterPro; IPR001569; HCV_NS5.
InterPro; IPR001909; HCV_CAPSId; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M58335; AAA72945.1; -. PIR; A38465; GNWVTC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PDB; 1CSJ; 08-NOV-99.
PDB; 1GX5; 09-APR-02.
PDB; 1GX6; 10-APR-02.
PDB; 1QUV; 26-JUN-00.
PDB; 80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB; 1NS3; 08-APR-98
PDB; 1C2P; 15-NOV-00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (RNA)(N).
SUBUNIT: T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
```

PF01006;

Pfam:

us-09-491-146a-32.rsp

```
81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D90208; BAA14233.1; -.
                                                                                           Local Similarity
hes 22; Conservative
                        1202
                                               1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISCUSSION OF SEQUENCE
 1187 1189
1189 1190
1198 1200
1203 1200
1680 1680
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A39253; GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11116;
                        HELIX
TURN
STRAND
SEQUENCE
                                                                                 Query Match
   TURN
                                                                                                                                                                                                 POLG_HCVJA
                                                                                                        Matches
                                                                                                                                                                                        RESULT
                                                                                                                                                                                                               51 F F F F S
                                                                                                                                ò
                                                                                                                                                    셤
                                                                                                               CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJOR ENVELOR (POTENTIAL).
NONSTRUCTURAL PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEBEZ/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLCNAC.
                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                     N-LINKED
N-L
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1.
BroDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                               645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       035
047
050
050
063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       085
087
092
094
097
103
1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1082
1086
1090
1093
                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                            ACT_SITE
                                                                                                                                                                                                                                 TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
STRAND
HELIX
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HELIX
```

```
D POLG_HCVJA STANDARD; PRT; 3010 AA.

C P26662;

O 1-40G-1992 (Rel. 23, Last sequence update)

T 01-40G-1992 (Rel. 23, Last sequence update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

T 28-FEB-2003 (Rel. 41, Last annotation update)

E Genome polyprotein [Contables: Capsid protein C (Core protein E E GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)

E EC 3 4-22.-); Protease/helicase NS3 (P70) (Hepacivirin)

E EC 3 4-22.-); Protease/helicase NS3 (P70) (Hepacivirin)

E EC 3 4-21.98); Nonstructural protein NS5A (P65); Nonstructural protein NS5B (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].

E Hepatitis C virus (isolate Japanese) (HCV).

C Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA.
                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEGUENCE FROM N.A.
MEDINE-9108550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Modecular cloning of the human hepatitis C virus genome from
"Appanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                         Score 121; DB 1; Length 3010;
Pred. No. 2e-09;
4; Mismatches 2; Indels
327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                 28
                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG
```

```
PKPKRQTKRNTLRRPKNVKFPAGGQIVG
                                  SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                   POLC_HCVH
                   -
                                            S
                                                                                                       POLG_HCVH
                                                                                        RESULT
                   à
                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                           Transferase; RNA-directed RNA polymerase;
W Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein C BY THE
CELLULAR AMINOPEPTIDASE.
THEMOVER C POTENTIAL;
CARIN 116 191 MATRIX PROTEIN (POTENTIAL).
THATIN 1007 1015 PROTEIN ROPEIN NS1 (POTENTIAL).
THAISMEN 1007 1015 PROTEINCHAL PROTEIN NS2 (POTENTIAL).
THANSMEN 1010 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
THANSMEN 3013 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
THANSMEN 3013 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
THANSMEN 3013 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
THANSMEN 3013 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DEPT (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 121; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED GLCNAC.
                                    InterPro; IPR001410; DEAD.
InterPro; IPR001522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001650; HCV_NS8.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001650; HCV_NS3.
InterPro; IPR01650; HCV_capsid; I.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01539; HCV_core; I.
Pfam; PF01506; HCV_NS1; I.
Pfam; PF01001; HCV_NS4; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2077
2240
2788
327017 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487; DEXDC;
             MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645
2041
2077
2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
```

Gaps

ö

2; Indels

Pred. No. 2e-09;

78.68;

Conservative

Best Local Similarity Matches 22; Conserv

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions also as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                    01-406-1992 (Rel. 23, Created)
01-406-1992 (Rel. 23, Last sequence update)
15-5EP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein [Contains motetation update)
67-003 (Rel. 20, Last annotation update)
67-004 (Rel. 20, Last annotation update)
67-005 (Rel. 20, Last annotation update)
67-006 (GP69) (GP10) (Rel. 20, Protein Pl. 20, Protein Protein Protein Protein NS2 (PP1)
67-006 (GP69) (Rel.) Protease/hellcase AS3 (PP0) (Hepacivirin)
6. C 3.4.21.98); Nonstructural protein NSA (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
6. Reattis C virus (isolate H) (RC) (Mana stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatisis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -i- FUNCTION: PROTEASE NS IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-1- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A, NS4B-NS5A, NS4B-NS5B, NS3-NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154121; PubMed-9493370;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RNA)(N).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-1- PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-1- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY 199.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prince A.M.; "Genomic strain H of hepatitis "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                   nchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
```

```
(POTENTIAL)
        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81.2%; Score 121; DB 1; Length 3011; 78.6%; Pred. No. 2e-09; 1ve 4; Mismatches 2; Indels
  623 623 N-LINKED (GLCNAC. ..) (POTENTI)
124 125 645 N-LINKED (GLCNAC. ..) (POTENTI)
125 1236 1236
1237 1238 1238
1251 1251 1251
1251 1252 1277
1262 1277 1282
1277 1282
1271 1282
1283 1285
1291 1295
1291 1295
1302 1301
1317 1319
1317 1319
1318 1349
1318 1340
1341 1341
1341 1341
1341 1341
1342 1348
1343 1349
1341 1341
1410 1420
1411 1410
1419 1420
1439 1439
1450 1480
1480 1480
1481 1488
1481 1488
1481 1488
1481 1488
1481 1488
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1481 1481
1591 1511
1514 1527
1550 1550
1570 1570
1584 1597
1584 1597
1584 1597
1587 1657
1687 1687
1687 1687
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
1681 1681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity 78.6
22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Si
Matches 22;
      CARBOHYD
CARBOHYD
CARBOHYD
TURN
TURN
TURN
TURN
TURN
STRAND
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                   ð
      (See http://www.1sb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DR Pfam; PF01543; HCV_corpsid; 1.

DR Pfam; PF01543; HCV_corpsid; 1.

DR Pfam; PF01543; HCV_corpsid; 1.

DR Pfam; PF01550; HCV_uns; 1.

DR Pfam; PF01550; HCV_uns; 1.

DR Pfam; PF01560; HCV_uns; 1.

DR Pfam; PF01500; HCV_uns; 1.

DR Pfam; PF01006; HCV_uns; 1.

DR Pfam; PF01506; HCV_uns; 1.

DR Pf01506; HCV_uns; 1.

DR Pfam; Pf01506; HCV_uns; 1.

DR Pfam; Pf01506; HCV_uns; 1.

DR Pf01506; HCV_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
NOMSTRUCTURAL PROTEIN NS4.
NONSTRUCTURAL PROTEIN NS48.
NONSTRUCTURAL PROTEIN NS58.
NONSTRUCTURAL PROTEIN NS58.
NONSTRUCTURAL PROTEIN NS58.
ONDSTRUCTURAL PROTEIN NS58.
ONDSTRUCTURAL PROTEIN NS58.
ONDSTRUCTURAL PROTEIN NS58.
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
ATP (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
entities requires a license agreement (or send an email to license@isb.sib.ch)
                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid.
InterPro; IPR002512; HCV_capsid.
InterPro; IPR002513; HCV_core.
InterPro; IPR002513; HCV_NS1.
InterPro; IPR004019; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001609; HCV_NS4.
InterPro; IPR001609; HCV_NS4.
InterPro; IPR002166; HCV_NS4.
InterPro; IPR001009; RN_PO_LDS_PS.
InterPro; IPR0071095; RNA_PO_LDS_PS.
InterPro; IPR007095; RNA_PO_LSS_PS.
InterPro; IPR007095; RNA_PO_LSS_PS.
                                                  EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB; 1HEI; 25-NOV-98. PDB; 1A1V; 16-FEB-99. PDB; 1A1N; 17-JUN-99. MEROPS; U39.001; -. TRANSFAC; T04155; -. INTEPTVO; IPRO01410; DEAD. INTEPTVO; IPRO014010; DEAD. INTEPTVO; IPRO012522; HCV_Cap. InterPro; IPR002521; HCV_cop. InterPro; IPR002521; HCV_cop. InterPro; IPR00140; HCV_cop. InterPro; IPR002521; HCV_cop.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1100

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

1000

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

Gaps

;;

```
DECOUGLE HOUSE

TO POLG HOUSE

TO POLG HOUSE

TO 1-AUG-1992 (Rel. 23, Created)

TO 1-AUG-1992 (Rel. 23, Last sequence update)

TO 1-AUG-1992 (Rel. 24, Last sequence update)

TO 1-AUG-1992 (Rel. 24, Last sequence update)

TO 1-AUG-1992 (Rel. 27, Last sequence u
                                                                                                             W Polygrotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Helicase; ATP-binding; TINIT_MET 1 1 18 EMMOVED FROM CAPSID PROTEIN C BY THE CELLULAR ANINOPEPTIDASE.

T CHAIN 116 191 MATRIX PROTEIN (POTEWITAL).

T CHAIN 192 383 MAJOR ENVELOPE PROTEIN NS! (POTEWITAL).

T CHAIN 194 733 NONSTRUCTURAL PROTEIN NS! (POTEWITAL).

T CHAIN 734 1010 PROTEASE/HELICASE NS3 (POTEWITAL).

T CHAIN 1619 PROTEIN NS2 (POTEWITAL).

T CHAIN 167 2017 NONSTRUCTURAL PROTEIN NS4 (POTEWITAL).

T ACT_SITE 1067 1087 CHARGE RELAY SYSTEM (BY SIMILARITY).

T ACT_SITE 1111 1111 CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 121; DB 1; Lengtn Ju
Pred. No. 2.1e-09;
2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1091 N-LINKED (GLCNAC. .) (FOUR 2038 N-LINKED (GLCNAC. .) (POT 2811 N-LINKED (GLCNAC. .) (POT AA; 329165 MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SW00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                               A MEDLINE-9204440: PubMed-1658196;
A MEDLINE-9204440: PubMed-1658196;
A MEDLINE-92044440: PubMed-1658196;
A Okamoto H., Okada S.-I., Suglyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
from a human carrier: comparison with reported isolates for conserved
and divergent regions.
T. J. Gen. Virol. 72.2697-2704(1991).
J. Gen. Virol. 72.2697-2704(1991).
C -1- FUNCTION: THE SMALL PROFIEIDS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl'
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
            (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002868; HCV_NSSa.
IPR001166; HCV_RdRP.
IPR001169; Helicase_C.
IPR007095; RNA_POL_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002522; HCV_capsid.
IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01543; HCV_capsid; 1.
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF02907
                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
```

+

ö

Gaps

ö

Pfam; Pfam; Pfam;

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein (Contains enotetino update)
67-003 (Rel. 42, Last annotation update)
67-004 (NSI); Protein P7; Nonstructural protein NS2 (PS2);
67-03 (GP08) (GP09) (NSI); Protease/hellcase NS3 (P70) (Hepacivirin)
6. C. 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (PS6); NONSTRUCTURAL NS5A (PS6);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PREDIEGRAPY, LIN M. H., Tail K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J., Lin M. H., Tail K.F., Liu p.C., Lin C.J., Chen D.S.;

The Taiwanese hepatitis of viral genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";

Virology 188:102-13(1992).

INCATION: THE SMALL PROYEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

I. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

I. CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
          RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                             SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                       (POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 121; DB 1; Length 3033;
Pred. No. 2.1e-09;
4; Mismatches 2; Indels (
                                                             (BY
(BY
(BY
                                                       CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                               N-LINED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC.
                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 2.1e
4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKROTKRNTLRRPKNVKFPAGGOIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-92230206; PubMed-1314449;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    330177 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVTW
ID POLG_HCVTW
AC P29846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
       CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                          ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLYPROTECTION DEALW.

POLYPROTECTION DEALW.

W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;

W Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

T INIT_MET 1 REWOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.

T CHAIN 1 115 CAPSID PROTEIN (POTENTIAL).

T CHAIN 192 383 MAJOR ENVELOPE PROTEIN NSI (POTENTIAL).

T CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1011 1619 PROTESSE NSI (POTENTIAL).

T CHAIN 1620 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
                                                                                                                                          SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SEQUENCE FROM N.A.
MEDLINE-92230232; Pubmed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                     *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS5.
InterPro; IPR001608; HCV_RGRP.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR001094; RNA_pol_DS_PS.
InterPro; IPR001509; HCV_capgid; 1.
Pfam; PF01542; HCV_capgid; 1.
Pfam; PF01518; HCV_NS1; 1.
Pfam; PF01509; HCV_NS2; 1.
Pfam; PF01506; HCV_NS3; 1.
Pfam; PF01506; HCV_NS3; 1.
Pfam; PF01606; HCV_NS3; 1.
Pfam; PF01606; HCV_NS3; 1.
Pfam; PF01609; HCV_NS3; 1.
Pfam; PF01609; HCV_NS3; 1.
Pfam; PF01609; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_core.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410;
InterPro; IPR00252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002519;
InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A40250; GNWVJB.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS
```

ö

Gaps

```
EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11104;
                                                                                                        623
645
2041
2077
2240
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N)
          CARBOHYD
CARBOHYD
                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCV1
                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                           CARBOHYD
                                                                                                          CARBOHYD
                                                                                                                       CARBOHYD
                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                    P26664;
                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
POLG_HCV1
           g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR ANINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI, PEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI, PEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI, PROTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
RAA-DIRECTED RNA POLYMERASE (POTENTIAL).
CHARGE RELAX SYSTEM (BY SIMILARITY).
CHARGE RELAX SYSTEM (BY SIMILARITY).
CHARGE RELAX SYSTEM (BY SIMILARITY).
ATT (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FIGURE, ELECTOR SMOO487; DEXDC; 1.
SMART; SMOO487; DEXDC; 1.
POlyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                      InterPro; IPR002222; HCV_capsid.
InterPro; IPR002212; HCV_capsid.
InterPro; IPR002513; HCV_core.
InterPro; IPR002513; HCV_core.
InterPro; IPR002513; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR007666; HCV_RGRD.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PS_IPR007094; RNA_POL_PS_IPR007094; RNA_POL_PS_IPR007094; RNA_POL_PS_IPR007094; HCV_RS_IPR007094; HCV_RS_IPR007094; HCV_RS_IPR007094; HCV_NS_IPR007094; HCV_NS_IPR007099; VIRAL_RGRP; IPR007099; VIR
                                                                                                                                                                                       EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                             PIR; A40244; GNWVTW.
PDB; 1N64; 25-FEB-03
PDB; 1N53; 08-APR-98
                                                                                                                                                                                                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

CARAMA DARAMA CARAMA CA

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseé(isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

MEDLINE—91172826; PubMed=1848704;

RA Callegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Medina—Selby A., Barr P.J., Weiner A.J.,

RA Gallegos C., Coit D., Weiner A.J.,

RA Callegos C., Coit D., Wein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                  (POTENTIAL)
      (POTENTIAL)
                                        POTENTIAL
                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC. . ) (POTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                   Score 114; DB 1; Len'
                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                           (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 KPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                             623
645
2041
2077
2529
2788
1; 327047 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                         76.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 76.5
Best Local Similarity 77.8
Matches 21, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                ¥
```

Gaps

```
STRAINCESTEL, GS. TISSUE-Embryo;

XA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

A Rawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

A Arawa J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

A Arawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

A Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

Radota K., Matsuda H.A., Sabburner M., Banalov S., Casawant T.,

Radota K., Matsuda H.A., Sabburner M., Baralov S., Casawant T.,

A Kuchi P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Ruchi P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Sakai K., Okido T., Furuno M., Ano H., Baldarelli R., Barsh G.,

Brownstein M.J. Bult C., Fletcher C., Fujita M., Gariboldi M.,

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

Sasaki H., Toyo-oka K., Hasegawa Y., Kawaji H., Kohtsuki S.,

Hayashizaki Y.;

Hayashizaki Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Liao C., Zhao M., Li T.;
"The expression of mouse LPTS1, a homolog of human tumor suppressor
LPTS, in mouse liver.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                  99C2X5; Q91W29; Q9D0C2; 28-FEB-2003 (Rel. 41, Created) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 41, Last sequence update) Pin2-interacting protein X1 (TRF1-interacting protein 1) (Liverrelated putative tumor suppressor) (LPTS1) (67-11-3 protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'Functional annotation of a full-length mouse cDNA collection.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE-21558139; PubMed-11701125;
Zhou X.Z., Lu K.P.;
"The Pin2/TRF1-interacting protein PinX1 is a potent telomerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 113; DB 1; Length 3011;
Pred: No. 3.1e-08;
5; Mismatches 3; Indels (
327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                          1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                           75.8%;
ilarity 71.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 38-332 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    speckles (By similarity)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nature 409:685-690(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cell 107:347-359(2001).
3011 AA;
                                                                  Local Similarity
nes 20; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IISSUE-Embryo;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  PINXI OR LPTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-ICR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           inhibitor
  SEQUENCE
                                              Query Match
                                                                       Best Loc
Matches
                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                         Interpro; IPR007094; RNA_POI_PS-I:

Pfam; PF01543; HCV_capsid; 1.

Pfam; PF01542; HCV_carv; 1.

Pfam; PF01542; HCV_carv; 1.

Pfam; PF01560; HCV_RNS; 1.

Pfam; PF01500; HCV_NS2; 1.

R Pfam; PF01000; HCV_NS3; 1.

R Pfam; PF01000; HCV_NS4; 1.

R Pfam; PF00271; HCV_NS4; 1.

R Probom; PD186082; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLUIAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/EZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
CHARGE RELAY SYSTEM (ECHARGE RELAY SYSTEM (FCHARGE RELAY SYSTEM (ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                     IPR001650; Helicase_C.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                 HCV_capsid.
                                                                                                                                                                                                                                                                                                       HCV_NS5a.
                                                                                                                                                                                                                                                                  HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                    IPRO02518;
IPRO04109;
IPRO00745;
IPRO01490;
                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001650;
InterPro; IPR007095;
                                                                                                                                 InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                              InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                            IPR002868;
                                                              MEROPS; S29.001;
                                                                                   U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3D-structure.
INIT_MET
                                                                                                                                                                                                                      InterPro; InterP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

```
Search completed: August 7, 2003, 11:20:05 Job time : 5.90909 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-8826585; Pubmed-2838958;
MEDLINE-8826585; Pubmed-2838958;
Collett M.S., Larson R., Belzer S.K., Retzel E.;
Collett M.S., Larson R., Belzer S.K., Retzel E.;
Collett M.S., Larson R., Belzer S.K., Retzel E.;
Ovroids encoded by bovine viral diarrhea virus: the genomic organization of a pestivirus.";
Virology 165:200-208(1988).
-1- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN WITH HELICASE AND PROTEASE ACTIVITY.
-1- PIM: GRILG GIVES RISE TO GP62 AND GP53; GP62 IN TURN YIELDS GP48
                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus.";
Virology 165:191-199(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Virusės; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bovine viral diarrhea virus (isolate NADL) (BVDV) (Mucosal disease
                                                                                                                                                                                                                              GO; GO:0005730; C:nucleolus; IDA.
GO; GO:0005730; C:nucleolus; IDA.
GO; GO:0005515; F:protein binding activity; IPI.
GO; GO:0007004; P:telomerase-dependent telomere maintenance; IDA.
InterPro: IPRO10467; G_Patch.
Pfam; PF01585; G-patch; 1.
PROSITE; PS50174; G_PATCH; 1.
Anti-onogenes; Nuclear protein; Chromosomal protein; Telomere.
DOMAIN
                                                                                                                                                                                                                                                                                                                                                              TELOMERASE INHIBITORY DOMAIN (TID)
ETPV -> HAPC (IN REF. 3);
687147CED58D991A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
 DOMAIN: The TID (telomerase inhibiting domain) domain is sufficient to bind TERT and inhibits its activity (By
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-88265858; PubMed-2838957;
Collett M.S., Larson R., Gold C., Strick D., Anderson D.K.,
Purchio A.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 332;
2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                            7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-FEB-1991 (Rel. 17, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3988 AA
                                                                                                                                                                                                                                                                                                                                                                                                                 34.9%; Score 52; DB 40.9%; Pred. No. 2.9; ive 6; Mismatches
                                                                                                                                                    EMBL; AF421879; AAL32445.1; -.
EMBL; AF221817; AAL37221.1; -.
EMBL; AK012057; -: NOT_ANNOTATED_CDS.
EMBL; AK011578; -: NOT_ANNOTATED_CDS.
EMBL; AJ344106; CAC51439.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |||||: : |:||: |
PKPKKRKAKKKLQRPEGVEIDA 316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPA 22
                                                                                                                                                                                                                                                                                                                                                                             210 213 E'
332 AA; 37221 MW;
                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 40...
Best Local 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                      MGI:1919650; Pinxl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11100;
                         similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_BVDVN
P19711:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pestivirus
                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     virus).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
POLG_BVDVN
ID POLG_BV
AC P19711;
DT 01-FEB-
DT 01-FEB-
                                                                                                                                                                                                                       MGD;
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Helicase; Serine protease; Hydrolase. CHAIN 1 7270 P20 (30KD). CHAIN 771 71063 GP115/GP62-GP53 (GLYCOPROTEIN). CHAIN 7 GP125/GP54-GP80.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 3988;
-i- SIMILARITY: TO THE HOG CHOLERA VIRUS GENOME POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4474212F338661BB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. 47;
5; Mismatches
                                                                                                                                                                                                                                                                     InterPro; IPR001280; CDV1r_endptseP80.
InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR00165; Helicase_C.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001695; RNA_POl_DS_PS.
InterPro; IPR001569; RNA_POl_DS_PS.
InterPro; IPR001569; RNASe_T2.
Pfam; PF00271; helicase_C; I.
Pfam; PF00998; Viral_RdRp; I.
PRNNS; PR00729; CDVENDOPTASE.
SMART; SM00497; DEADC; I.
SMART; SM00490; HELICC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 51.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3189 PKPKQVTKEAVRNLIEQKKDVEIP 3212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTK---RNTLRRPKNVKFP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   449154 MW;
                                                                                                                                                                                        EMBL; M31182; AAA42854.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     34.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity 45.8
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ä
                                                                                                                                                                                                          PIR; A29198; GNWVBV
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                        MEROPS; C53.001; -. MEROPS; S31.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1451
1803
2224
2307
2772
2772
2981
3778
3867
3988
```

us-09-491-146a-32.rspt

```
hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                               hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                             hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                 hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis
hepatitis
                                                                                                                                                                                                                                                                                              hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                  hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chayama K.;
Submitted (FEB-1955) to the EMBL/GenBank/DDBJ databases.
Submitted (FEB-1955) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D49374; BAAD0 MRNA (BY SIMILARITY).
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus type 3b.
Hepatitis C virus type 3b.
Hepatitis C virus type 3b.
 081233
081231
081234
081234
081234
068861
068863
068863
068863
068955
068955
068951
068308
068308
0683108
0683108
0683108
068310
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
0683108
06831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3023 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
 081233
0812334
0812334
0812346
081234
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068865
068965
068965
068967
068967
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
068067
                                                                                                                                                                                                                                                                                                                                                                                                                              080P84
080P75
080P79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
PE01543; HVV_capsid; 1.
Pfam; PP01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
 1009
11099
11099
11099
11099
11099
11099
                                                                                                                                                                                                                                                                                                                               000
                                                                                                                                                                                                                                                                                                                                                               2222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519;
InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR004109;
InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
 NCBI_TaxID-42791;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001;
MEROPS; U39.001;
081487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                081487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1
 081487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q81487 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hepatitis
                                                                                                         August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         081550
081772
084782
091725
091725
091726
068413
081266
08126
081230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P87760
P87761
                                                                                                                                                                                                                                                                                                                             830525
                              Compugen Ltd.
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compug
                                                                                                                                                                                                                                                                                                                        tal number of hits satisfying chosen parameters:
                                                                                                                                                                                         149
1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                           830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P87760

P87761

Q81550

Q8QYR2

Q91KZ5

Q91KZ7

Q91KZ6

Q68413

Q81266

Q68305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08JYR3
081230
081235
                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_fungi:*
sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                          seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_archea:*
sp_bacteria:*
                                                                                                                                                                            US-09-491-146A-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_rvirus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           419
108
108
415
415
100
191
191
191
43
43
109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.9
87.2
87.2
87.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0
                                                                                                                                                                                             Perfect score:
                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB s
Maximum DB s
                                                                              OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database :
                                                                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                                           Searched:
                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1111110987654
1111110987654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ş
```

Gaps

ö

```
Pfam; PF01560; HCV_NS1; 1.

PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.

Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.

CHAIN 1 >190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-96305314; PubMed-8712927; Ohloo E., Ohba K.-I., Wu R.-R., Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba K.-I., Wu R.-R., Kolde T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R.; "Usefulness and limitation of phylogenetic analysis for hepetitis C virus core region: application to isolates from Egiptian and Yemeni
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Arch. Virol. 141:1101-1113(1996).
-1- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
-1- LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ohno T., Mizokami M.;
"Determination of nine genotypes of hepatitis C virus using PCR
method.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 88.6%; Score 132; DB 12; Length 108; Best Local Similarity 85.7%; Pred. No. 2.1e-12; Matches 24; Conservative 2; Mismatches 2; Indels (
                                                                                                                                                                                                                                           Length 419;
                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                   191 >419 ENVELOPE PROTEIN.
419 419
419 AA; 45391 MW; CB8956E32EA6DD6C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER 108 108
SEQUENCE 108 AA; 11773 MW; A05A3836EC14BA21 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Thesis (1995), Nagoya City University Medical School
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P87761 PRELIMINARY; PRT; 108 AA.
P87761;
01-MAY-1997 (TrEMBLrel. 03, Created)
01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P87760 PRELIMINARY; PRT; 108 AA. P87760; Ol-MAY-1997 (TrEMBLrel. 03, Careated) 01-MAY-1997 (TrEMBLrel. 03, Last sequence update) 01-MAY-1997 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                        Score 137; DB 12;
Pred. No. 1.6e-12;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRQTKRNTTRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   patients.";
                                                                                                                                   CHAIN
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
P87760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P87761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               **Numcade H.;*
**Numcade H.;*
**Numcade H.;*
**Numcade H.;*
**Numcade H.;*
**Inucleotide sequence hepatitis C virus (type 3b) isolated from a Japanese patient with chronic hepatitic C.";

**Inucleotide sequence hepatitis C.";

**Inucleotide sequence he
                                                                                                                                                                                                                  ProDom; PD186062; HCV_NS1; 1.
SWART; SMO0487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
CHAIN 191 384 E1.
CHAIN 192 384 E1.
CHAIN 385 737 E2/NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-94270990; Pubmed-7545932;
Hotta H., Doi H., Hayashi T., Purwanta M., Soemarto W., Mizokami M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         firuses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Analysis of the core and El envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia."; Arch. Virol. 136:53-62(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-95086611; Pubmed-7996156;
Chayama K., Tsubota A., Koida I., Arase Y., Saitoh S., Ikeda K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 149; DB 12; Length 3023; 100.0%; Pred. No. 2e-13; 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     329734 MW; 5268D1EC410AC545 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
101-NAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           419 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                    Pfam; PF01006; HCV_NG4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF01506; HCV_NS5a; 1.
Propon; PF0666; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 100. Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1014
1665
1980
3023
                                                                                                                                                                                                                                                                                                                                                                                                        192 38
385 73
738 103
1015 166
1666 198
1981 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ohba K., Homma M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                          PF01538; PF02907; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        081345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
        Pfam;
                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
```

à

ö

Gaps

```
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
SEQUENCE
                                                                                                                                      SEQUENCE
                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                         O8JYR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8QP82
                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 7
                                                                                                                                                                                                                                                                 RESULT
Q8JYR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   080P82
 STEERS
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OC OE DIT
                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BY A
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                              Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba K.-I., Wu R.-R., Kolde T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R.; "Usefulness and limitation of phylogenetic analysis for hepatitis C virus core region: application to isolates from Egiptian and Yemeni
   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                    FACE. VICOL. 141:1101-1113(1996).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A ND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).

EMBL, D82033; BAA11518.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                 Ohno T., Mizokami M.; "Determination of nine genotypes of hepatitis C virus using PCR
                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M., Tanaka T., Miyakawa Y., Mayumi M.; "Hepatitis B virus subtypes and hepatitis C virus genotypes in patients with chronic liver disease in Nepal."; Hepatology 19:805-809(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Core, env, and part of E2/NS1 (Genome polyprotein) (Fragment). Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                               Length 108;
                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                         11773 MW; A05A3836EC14BA21 CRC64;
                                                                                                  Thesis (1995), Nagoya City University Medical School
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (TrEMBLrel. 01, Last sequence update) (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                              Score 132; DB 12;
Pred. No. 2.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           415 AA
                                                                                                                                                                                                                                                                                                                                                                     2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRQTKRNTTRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                            1 PKPKROTKRNTLRRPKNVKFPAGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                    MEDLINE-96305314; PubMed-8712927;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-NE137;
MEDLINE-94186155; PubMed-8138250;
                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV_caps1d.
Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                              88.68;
                                                                                                                                                                                                                                                                                                                                                                     24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01-OCT-2002 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                         108 AA;
                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                        SEQUENCE FROM N.A.
Hepatitis C virus
                                                      SEQUENCE FROM N.A
                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                  Polyprotein.
NON_TER 1
SEQUENCE 1(
                                                                                                                                                                                               patients.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081550
                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 5
081550
ID 08151
AC 08151
DT 01-N
DT 01-N
DE COTE
OS HEPA
                                                                                                                                                                                                                                                   CC
DR KW
DR CC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
EMBL; D16616; BAA04038.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01560; HCV_NS1; 1.
Polyprotein; Interproper protein; Glycoprotein; Nonstructural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ⋖
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepartitis C virus in Western Siberia.";
"Genetic variability of hepartitis C virus in Mestern Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                   Length 415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                            OAD63410FA8F522F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             73 AA; 8272 MW; 64D60827A48FD796 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-2002 (TrEMBLrel, 22, Created)
01-0CT-2002 (TrEMBLrel, 22, Last sequence update)
01-MAR-2003 (TrEMBLrel, 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C_virus.
                                                                                                                                                                                                                                                                                                                                                                             88.6%; Score 132; DB 12;
85.7%; Pred. No. 9.1e-12;
ive 2; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87.9%; Score 131; DB 12;
82.1%; Pred. No. 1.9e-12;
ive 4; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              73 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQNVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF506623; AAM33399.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein (Fragment). Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                            415 AA; 44703 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 82.10,
                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 85.7 nes 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-RIG304;
```

191 AA

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
es 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                Hepatitis C virus
                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                         STRAIN-SOM1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-SOM2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091K26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          091K26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
Q91K26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
     g
                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                     Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";

J. Virol. 76:4034-403(2002).

-! SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

INTERPRO: IPR002522; HCV_capsid.

Petan; PF01543; HCV_capsid, 1.
                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-2003;
MEDLINE-2073994; PubMed-10811932;
Abid K., Quadri R., Veuthey A.L., Hadengue A., Negro F.;
And K., Quadri R., Veuthey A.L., Hadengue A., Negro F.;
A novel hepatitis C virus (HCV) subtype from Somalia and its
classification in HCV clade 3.";
J. Gen. VIROL 81.1485-1493 (2000)
I. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                             uery Match 87.9%; Score 131; DB 12; Length 100; est Local Similarity 82.1%; Pred. No. 2.7e-12; Matches 23; Conservative 4; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.9%; Score 131; DB 12; Length 191; 82.1%; Pred. No. 5.5e-12; 1ve 4; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
Core structural protein (Genome Polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                          Polyprotein.
NOW_TER 100 100
SEQUENCE 100 AA; 11235 MW; F2A0FD95F9E85AEE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20730 MW; 7E85DDDE0DC4AC3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 15, Created)
(TrEMBLrel. 15, Last sequ
(TrEMBLrel. 22, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                      MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF216794; AAF44741.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 82.1
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                    FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11103;
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9IK25;
01-OCT-2000 (
01-OCT-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     091K25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
```

**Q91K25** 

ŝ g

ð a

```
Abid K., Quadri R., Vetthey A.L., Hadengue A., Negro F.;

"A novel hepatitis C virus (HCV) subtype from Somalia and its
"T classification in HCV clade 3.";
"J. Gen. Virol. 81:1485-1493(2000).

"I. Guburi THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BYNELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

"EMBL; AF216792; AAF44739.1; -...

"R InterPro; IRN005221; HCV_capsid.

"InterPro; IRN005221; HCV_capsid.

"InterPro; IRN005221; HCV_capsid.

"InterPro; IRN005221; HCV_capsid.

"InterPro; IRN005221; HCV_capsid."

"InterPro; IRN005221; HCV_capsid."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Abid K., Ouddri R., Veuther A.E., Hadengue A., Negro F.;
"A novel hepatitis C virus (HCV) subtype from Somalia and its
classification in HCV clade 3.";
J. Gen. Virol. 81:1485-1493(2000)
-1- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AF216793; AAF44740.1;
                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 191;
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core structural protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-007-2000 (TrEMBLrel. 15, Created)
01-007-2000 (TrEMBLrel. 15, Last sequence update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
Core structural protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191 191
191 AA; 20730 MW; 7E85DDDE0DC4AC3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20730 MW; FC844F8E1573F615 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 131; DB 12;
Pred. No. 5.5e-12;
4; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PKPQRKTKRNTIRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=20273994; PubMed=10811932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-20273994; PubMed-10811932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.9%;
ilarity 82.1%;
Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SO TWENT TO BE THE SO OCCOSE T
```

```
-i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                              InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01542; HCV_capsid; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_caps1d. Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87.2%;
illarity 82.1%;
Conservative 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 22, (TrEMBLrel. 22, 1 (TremBLrel. 23, 1
                                                                                                                                                                                                                                    Match 87.9%;
Local Similarity 85.7%;
les 24; Conservative
                                                                                                                                             Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8JYR3 PRELIMINARY;
Q8JYR3;
01-OCT-2002 (TFEMBLREL: 2:
01-CCT-2002 (TFEMBLREL: 2:
01-MAR-2003 (TFEMBLREL: 2:
                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
es 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-HCV-BB46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                   068305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q8JYR3
                                                                                                                                                                                                                                                                                                                                                                      068305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  44441
                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                             ò
     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             five novel
9c) major
                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                   ø
                                                                                                                                                                                                                                                                                                       MEDLINE-97052554; PubMed-8897188; Bernier L., Willems B., Delage G., Murphy D.G.; Identification of numerous hepatitis C virus genotypes in Montreal, Canada.";
                                                                                                                                                                                                                                                                                                                                                                J. Clin., Microbiol. 34:2815-2818(1996).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (8Y SIMILARITY).

EMBL, 033437; AAB40040.1; -.

InterPro: IPR002522; HCV_capsid.

InterPro: IPR002521; HCV_core.

Pfam; PF01543; HCV_core.

Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                               ö
 Score 131; DB 12; Length 191;
Pred. No. 5.5e-12;
4; Mismatches 1; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tokita H., Okamoto H., Luengrojanakul P., Vareesangthip K., Chainuvati T., Iizuka H., Tsuda F., Miyakawa Y., Mayumi M.; "Hepatitis C virus variants from Thailand classifiable into genetypes in the sixth (6b), seventh (7c,7d) and ninth (9b,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87.9%; Score 131; DB 12; Length 191; 82.1%; Pred. No. 5.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191
20720 MW; 7A4B84532A1D07F9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  415 AA.
                                                                                                                                                 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4; Mismatches
                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                      5 PKPQRKTKRNTIRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genetic groups.";
J. Gen. Virol. 76:2329-2335(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Th576;
MEDLINE-96005057; PubMed-7561773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein (Fragment).
Hepatitis C virus.
      87.9%;
82.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 82.1 ies 23; Conservative
                 Best Local Similarity 82.1
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  081266
                                                                                                                                                968413
                                                                                                                     RESULT 11
Q68413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                g
                                                                                                                                                             à
```

```
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Songsivilal S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis {\tt C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
NON_TER 415 415 SEQUENCE 415 A3, 44881 MW; FD151EF15B1C31FC CRC64:
                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDEJ databases.
Submitted Submitted of THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GEXCOPROTEIN B. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, U23743; AAA65050 1;
                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                          Length 415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;; Score 130; DB 12; Length 43;
;; Pred, No. 1.5e-12;
·4; Mismatches 1; Indels
                                                                                                                                                                                            2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 43 43 43
43 AA; 4844 MW; Al73B6263AF6DC4F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                          Score 131; DB 12;
Pred. No. 1.3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                           43 AA.
                                                                                                                                                                                            2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                      PRT;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                             SEQUENCE FROM N.A.
STRAIN-RIG292;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPEID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; AF506622; AAM33398.1;
-InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Office T. Mizokami M., Chstopher T., Obba K., Suzuki K., Roger W.;
"The nucleotide sequence of the core region of HCV in Pakistan and
Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
--- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; D29647; BAA21022.1;
PIGHEPFO; IPRO02522; HCV_capsid.
PÉQUISA3; HCV_capsid; 1.
                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.2%; Score 130; DB 12; Length 109; 85.7%; Pred. No. 4.3e-12; 1.ve 2; Mismatches 2; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 130; DB 12; Length 73;
Pred. No. 2.8e-12;
4; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER 109 109
SEQUENCE 109 AA; 12293 MW; 524E8D425A7E3CB2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76549827A490C896 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              109 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       73 73
73 AA; 8258 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 87.2%;
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Ban1-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        081230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
081230
10 00123
AC 08123
AC 08123
AC 08123
DT 01-NO
DT 01-NO
DC COFE
OS VITUS
OC VITUS
OC HEPAC
OC HE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

Search completed: August 7, 2003, 11:19:03 Job time : 26.6364 secs

us-09-491-146a-32.rai

Sequence 32, Application US/08921887

Patent No. 603071

GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP Sequence 3, Appli Sequence 26, Appli Sequence 9, Appli Sequence 23, Appli Sequence 3, Appli Sequence 10, Appli Sequence 14, Appl Sequence 8, Appli Sequence 8, Appli Sequence 8, Appli Sequence 8, Appli Sequence 2, Appli Sequence 2, Appli 60, Appl 1, Appli 2, Appli Sequence 10, Appl Sequence 198, App Sequence 232, App Gaps Sequence 46, Sequence 60, Sequence ö Length 28 Indels COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWARE: PatentIn Release #1.0, Version #1.30
APPLICATION NUMBER: US/08/921,887 ADDRESSEE: JONES & ASKEW, LLP STREET: 191 Peachtree Street, N.W., 37th Floor CITY: Atlanta 100.0%; Score 149; DB 3; 100.0%; Pred. No. 3.3e~15; tive 0; Mismatches 0; US-08-530-550-3 US-08-262-037-26 PCT-US94-07088-9 PCT-US94-07088-9 US-08-836-075A-10 US-08-835-86C-198 US-08-835-86C-232 US-08-835-86C-232 US-08-835-86C-232 US-08-344-616-8 US-08-344-616-8 US-08-344-616-8 US-08-34-977-8 US-08-36-075A-46 US-08-836-075A-46 US-08-836-075A-46 US-08-836-075A-60 US-08-836-075A-60 ALIGNMENTS ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFESENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700 TYPE: amino acid STRANDEDNESS: not relevant TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHETICAL: NO TELEFAX: 404-818-3/99
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids ORGANISM: Hepatitis virus Query Match 100. Best Local Similarity 100. Matches 28; Conservative internal 61 61 61 61 61 74 74 74 74 74 74 74 74 74 74 74 74 FILING DATE: CLASSIFICATION: 435 CIT:. GA STATE: GA COUNTRY: USA TO: 30303-1769 ANTI-SENSE: NO FRAGMENT TYPE: US-08-921-887-32 US-08-921-887-32 191, App 192, App 175, App 175, App 6, Appli 36, Appli Sequence 3 Sequence Sequence Sequence 1 Sequence Sequence Sequence Sequence

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE OF INVENTION: NUCLECTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: AMINO ACID SEQUENCES OF HEBRATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 126; DB 2; Length 191;
Pred. No. 5.8e-11;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                           Length 28;
                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                    Score 126; DB 3; L
Pred. No. 7.3e-12;
5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                               1 PKPKROTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RECISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMONICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 187, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELETAN 1792
TELEX 421792
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
TYPE: amino acid
STRANDEDNESS: unknown
                                     LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
23:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              84.6%;
illarity 78.6%;
Conservative 5
                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative 5
                                                                                                                                                                                                                     ; ORGANISM: Hepatitis virus US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: MORGAN & FINI
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                           ANTI-SENSE: NO ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOMARD A.
TITLE OF INVENTION: MOSALC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                           APPLICANT: LEROUX-ROELS, Geert
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MARTENS, Geert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR PILING DATE: 1993-11-04
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 137; DB 4; Length 319;
Pred. No. 2.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                            PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 435
ATTORNEY_AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                      ; Sequence 211, Application US/08635886C
; Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 23, Application US/08921887 Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: Atlanta
STATE: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                   US-08-635-886C-211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-635-886C-211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 211
LENGTH: 319
                     ò
                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
```

ò

```
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS: ADDRESS: ADDRESSE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK STREET: NEW YORK COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE OF INVENTION: MUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84.6%; Score 126; DB 2; 78.6%; Pred. No. 5.8e-11; Live 5; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                          ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: PCT/US95/10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 187, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SYSTEM: PC-DOS/MS-DOS WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION UNBRER: 36,459
REFERENCE/DOCKET NUMBER: 2026-
TELEPHONE: (212) 758-4800
TELEPHONE: (212) 751-6849
TELEFX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 190: SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acids
TYPE: amino acids
TYPE: amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: NEW YORK
STATE: NEW YORK
COUTRY: USA
2IP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
COMPUTER: OF CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ) ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK12
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 78.6'
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                       NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HERATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84.6%; Score 126; DB 2; Length 191; 78.6%; Pred. No. 5.8e-11; Live 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
CORRESPONDENCE ADDRESS:
                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                   SOFTWARE:

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

IS APPLICATION NUMBER:

CLASSIFICATION:

ATS

ATONNEY/ACENT INFORMATION:

NAME: RICHARD W. BORK, 459

REFERENCE/DOCKET NUMBER: 2026-4116

TELECOMMUNICATION INFORMATION:

TELECHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEFAX: (212) 751-6849

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 188:

SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             111:1:11111:111:1111 | 11111 | 5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 190, Application US/08290665A
Fatent No. 5882852
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND TITLE OF INVENTION: NUCLECTIDE AND DE TITLE OF INVENTION: AMINO ACID SEQUENCE INVENTION: CORE GENES OF ISS
                                                                                                                                                                                                                                                                       Sequence 188, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 84.67
Best Local Similarity 78.67
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGINAL SOURCE
                                                                                                                                                                                                                               US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      . a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

Gaps

```
PCT-US95-10398-190

FCT-US95-10398-190

Sequence 190, Application PC/TUS9510398

GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: BURNTAION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263

CORRESPONDENCE ADDRESS:
ADDRESSE:
ADDRES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84.6%; Score 126; DB 5; Length 191; llarity 78.6%; Pred. No. 5.8e-11; Conservative 5; Mismatches 1: Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. 5.8e
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUFTWARE: WONDERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION DATA:
PRIOR APPLICATION UNBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION UNBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PC-DOS/MS-DOS
                                                                                 TELERAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDENESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 202
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEX: 421792,
INFORMATION FOR SEQ ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                 TELEFAX: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              T: 345 PARK AVENUE
NEW YORK
S: NEW YORK
RY: USA
                                                                                                                                                                                                                                                                                                                                                INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                  unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                               TOPOLOGY: unk
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUCLECTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 84.6%; Score 126; DB 5; Length 191; Best Local Similarity 78.6%; Pred. No. 5.8e-11; Matches 22; Conservative 5; Mismatches 1; Indels
FILING DAIE:
CLASSIFICATION
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
RECISTRATION NUMBER: 2026-4116
FELEPANION NUMBER: 2026-4116
FELECOMMUNICATION INFORMATION:
FLEEN: (212) 751-6849
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 188, Application PC/TUS9510398
SENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: BUKH, J., MILLER, R.H. TITLE OF INVENTION: NUCLEOTIDE AND DEDUC:
TITLE OF INVENTION: CORE GENES OF ISOLATION: TITLE OF INVENTION: AMINO ACID SEQUENCES:
TITLE OF INVENTION: AMD THE USE OF REAGES:
TITLE OF INVENTION: SEQUENCES IN DIAGNOS;
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE CITY: NEW YORK STAFF: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REGISTRATION NUMBER: 36,459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ....antSM: homosapiens
; INDIVIDUAL ISOLATE: HK10
PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
APPLICANT: DELEVS, SOURCE.

APPLICANT: DELEVS, Gert
APPLICANT: DELEVS, Gert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: PCT/FP94/03555
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR PILING DATE: 1994-11-04
NUMBER OF SEC ID NOS: 286
SOFTWARE: Patentin version 3.1
SEQ ID NO 230
: LENGTH: 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 126; DB 4; Length 319;
Pred. No. 1e-10;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; LOCATION: (233]..(233)
; OTHER INFORMATION: Xaa is any amino acid
US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (144)..(144)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME/KEY: MISC_FEATURE
LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ION: (161)..(161)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ON: (167)..(167)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ION: (171)..(172)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ION: (174)..(174)
INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: (232)..(232)
OTHER INFORMATION: Xaa is any amino acid
                   GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ch 84.6%;
1 Similarity 78.6%;
22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCATION: (149)..(149)
OTHER INFORMATION: Xaa 1s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
Matches 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                    RESULT to US-08836-075A-44

US-08-836-075A-44

Sequence 44, Application US/08836075A

Patent No. 6180768

GENERAL INFORMATION:
APPLICANT: MAETTENS, GEERT
APPLICANT: APPLICANT: LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES

TITLE OF INVENTION: AGENTS

TITLE OF INVENTION: AGENTS

TITLE OF INVENTION: AGENTS

TITLE OF INVENTION: AGENTS

TITLE OF INVENTION: AGENTS
                                                                                                                                                 ö
                                                                                                                                                 Gaps
                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                     Score 126; DB 5; Length 191;
Pred. No. 5.8e-11;
5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 3; Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA

CONFUTER TEADABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
SOFTWARE: M.Crosoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 126; DB 3;
Pred. No. 1e-10;
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/836,075A FILING DATE: 21 Apr 1997 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: EP 95870076.7 FILING DATE: 28 Jun 1995 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 230, Application US/08635886C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84.68;
78.68;
                                                                                                     Ouery Match
Best Local Similarity 78.6%;
Matches 22; Conservative
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK12
PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 78.65
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                          ö
                                                                                                                                                                                                                                  g
```

Gaps

```
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEBATITIES C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 125; DB 3; Length 74;
Pred. No. 2.9e-11;
3; Mismatches 2; Indels
                       COMPUTER: TEMP C COMPATIBLE
COMPUTER: TEMP C COMPATIBLE
SOFTWARE: TEMP C COMPATIBLE
SOFTWARE: MICROSOft WORD 6.0 / ASCII text OUTPUT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
FILING DATE: 21 Apr 1997
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 JUN 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 189, Application US/08290665A
Patent No. 5802852
GENERAL INFORMATION:
APPLICANT: BUKK, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILLIO DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                INNS: 004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 203
REFERENCE/DOCKET NUMBER: 203
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775 REFERENCE/DOCKET NUMBER: IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83.9%;
82.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE OF INVENTION: WILLDON TITLE OF INVENTION: OORE GE TITLE OF INVENTION: CORE GE TITLE OF INVENTION: SEQUENCE NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS: ADDRESS: ADDRESSE: MORGAN & FINNER STREET: AS AS PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : (212) 758-4800
(212) 751-6849
              Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS:
LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 82.1'
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-836-075A-104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                           APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 104, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS, C VIRUS GENOTYPES
TITLE OF INVENTION: AGENTS
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 83.9%; Score 125; DB 3; Length 28; Best Local Similarity 82.1%; Pred. No. 1e-11; Matches 23; Conservative 3; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 24, Application US/08921887 Patent No. 6030771 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
CLASSIFICATION: 435
ATGNREY AGENT INPERMATION:
NAME: WARREN, WILLIAM I.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3799
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TTYPE: amino acids
TTYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Hepatitis virus US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N: USA
77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                CITY: Atlanta
STATE: GA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-836-075A-104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: H
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
ò
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N.H.

NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                            Score 125; DB 2; Length 191;
Pred. No. 8.1e-11;
6; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PUNCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC
TITLE OF INVENTION: AMINO ACID SEQUENCES
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: AND THE USE OF REAGE
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:

NAME: RIGHARD W. BORK

REGISTRATION NUMBER: 36,459
REFERNCE/DOCKET NUMBER: 2026-4116
TELEPHONE: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 189, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MORGAN & FINNEGAN
TELEX: 421792
INFORMATION FOR SEQ 1D NO: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: 421792
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                              Query Match 83.9%;
Best Local Similarity 75.0%;
Matches 21; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET: 345 PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INDIVIDUAL ISOLATE: S2 PCT-US95-10398-189
                                                                                                                                                              ORGANISM: homosaplens
INDIVIDUAL ISOLATE: S2
US-08-290-665A-189
                                                                                TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: NEW YORK COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                            PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                          q
```

Appl

Sequence

Sequence Seq

OM protein

Run on:

Sequence:

Title:

```
APPLICANT: MAERTENS, GEERT
STUTYER, LIEVENS OF HEBATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES OF HEBATITIS C, THERAPEUTIC AND DIAGNO
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                      Sequence 23, Appl
Sequence 40, Appl
Sequence 4, Appl1
Sequence 9, Appl1
Sequence 1, Appl1
Sequence 20, Appl
Sequence 3, Appl1
Sequence 1, Appl1
Sequence 20, Appl1
Sequence 20, Appl1
Sequence 20, Appl1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPTR: 1203
COMPTR: 1203
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWARE: Microsoft Word 6.0 / ASCII text output
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 09/836,075
FILING DATE: -CURKNOWN-
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 OCt 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
US-09-878-281-42
US-09-878-281-44
US-09-878-281-44
US-09-19-95-2
US-09-19-138-18
US-09-851-138-18
US-09-851-138-18
US-09-89-046-50
US-09-899-046-50
US-09-899-046-51
US-09-899-046-14
US-09-878-281-50
US-09-878-281-50
US-09-973-025-50
US-09-95-808-50
US-09-91-611-23
US-09-95-808-50
US-09-91-611-23
                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-952-572-9
US-09-929-955-1
US-09-747-419-20
US-09-891-894-3
US-10-104-966-1
US-10-259-275-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
RANDLD, WHITE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 44, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES:
                                                                                            8318
3319
3319
3319
3319
8009
8009
                                                                                                                                                                                                                                                                                                                                                                       2894
2894
22985
3011
3011
3011
3011
3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-851-138-44
   Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 10, A Sequence 10, A Sequence 10, A Sequence 79, A Sequence 78, A Sequence 14, A Sequence 60, A Sequence 152, A Sequence 152, A Sequence 152, A Sequence 152, A Sequence 42, Ap Sequence 44, Ap
                                                                                                                         August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 104,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. (cgn2_6/ptodata///pubpaa/US07_PUBCOMB.pep:*
2. (cgn2_6/ptodata///pubpaa/US07_PUBCOMB.pep:*
3. (cgn2_6/ptodata///pubpaa/US06_NEW_PUB.pep:*
3. (cgn2_6/ptodata///pubpaa/US06_NEW_PUB.pep:*
3. (cgn2_6/ptodata///pubpaa/US07_NEW_PUB.pep:*
3. (cgn2_6/ptodata///pubpaa/US08_NEW_PUB.pep:*
3. (cgn2_6/ptodata///pubpaa/US08_NEW_PUB.pep:*
3. (cgn2_6/ptodata///pubpaa/US08_NEW_PUB.pep:*
3. (cgn2_6/ptodata///pubpaa/US09_PUBCOMB.pep:*
4. (cgn2_6/ptodata///pubpaa/US09_PUBCOMB.pep:*
4. (cgn2_6/ptodata///pubpaa/US09_PUBCOMB.pep:*
4. (cgn2_6/ptodata///pubpaa/US09_NEW_PUB.pep:*
4. (cgn2_6/ptodata///pubpaa/US10A_PUBCOMB.pep:*
4. (cgn2_6/ptodata///pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata///pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata///pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata///pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata//pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata//pubpaa/US10A_PUB.pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata//pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata//pubpaa/US10A_PUB.pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata//pubpaa/US10A_PUB.pep:*
4. (cgn2_6/ptodata//pubpaa/U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                       451899
                   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-851-138-44

US-09-851-138-104

US-09-851-138-104

US-09-758-308-1

US-09-758-308-1

US-09-921-397-77

US-09-921-397-77

US-09-921-397-78

US-09-851-138-46

US-09-851-138-46

US-09-851-138-46

US-09-851-138-60

US-09-891-138-14

US-09-891-138-14

US-09-891-138-14

US-09-899-046-152

US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                   .al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                  451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                            protein search, using sw model
                                                                                                                                                                                                                                                                                              Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                       length: 0
length: 2000000000
                                                                                                                                                                                                   US-09-491-146A-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  74
444
744
91
103
1113
1137
1166
1166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                              BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sed
                                                                                                                                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score
                                                                                                                                                                                                                       Perfect score:
```

Minimum DB Maximum DB

Searched:

Database :

Result

g

```
STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                            APPLICANT: PENIN, FRANCOIS
APPLICANT: DALBON, PASCAL
APPLICANT: LADISHON, PASCAL
APPLICANT: LADISHON, PASCAL
APPLICANT: LADISHON, PASCAL
APPLICANT: LADISHON, PASCAL
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT PELING DATE: 2003-02-19
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: BARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATENTING DATE: PATENTING DATE: 1998-03-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESSE:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
CORPUTER: MICROSOFT WORD
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B1.2%; Score 121; DB 15; 78.6%; Pred, No. 3.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/09/851,138 FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 08/836,075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL: Proc. Natl. Acad. Sci. U.S.A. VOLUME: 88 PAGES: 3392-3396 DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                          Sequence 1, Application US/10367677
Publication No. US20030118604A1
GENERAL INFORMATION:
APPLICANT: JOLIVET, MICHEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 78.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                     Query Match

84.6%; Score 126; DB 10; Length 319;
Best Local Similarity 78.6%; Pred. No. 5.5e-10;
Matches 22; Conservative 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         83.9%; Score 125; DB 10;
82.1%; Pred. No. 1.6e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
RADBRESSES: RNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 OCT 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US/09/851,138 FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-851-138-104
LENGTH: 319 amino acids
TYPE: amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-851-138-44
                                                                                                                                                                                                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                         1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                         : Sequence 104, Application US/09851138 ublication No. US20020183508Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 74 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION FOR SEQ ID NO: 104: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 82.1
Matches '23, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-851-138-104
                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81.2%; Score 121; DB 10; Length 97; 78.6%; Pred. No. 7.8e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 121; DB 10;
Pred. No. 8.3e-10;
4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GREERAL INCORNATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: SID nucleic acids and poly
TITLE OF INVENTION: pathogenic strain of the
TITLE OF INVENTION: applications thereof
FILE REFREENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                             CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 0S 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 14, Application US/09851138; Publication No. US20020183508A1 GENEAL INFORMATION: APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 77, Application US/09921397
Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                  TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 78.6%
                                                                                                                                                                                                                                                                                                                                                                                                                                                              97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 8
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TILLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TILLE REFERENCE: 14114.034912
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT FILING DATE: 2001-01-10
                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 81.2%; Score 121; DB 10; Length 74; Best Local Similarity 78.6%; Pred. No. 5.8e-10; latches 22; Conservative 4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
CITY: Washington
STATE: D. C.
COUNTRY: U. S.
ZIP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nery Match 81.2%; Score 121; DB 9; Jest Local Similarity 78.6%; Pred. No. 7.3e-10; Matches 22; Conservative 4; Mismatches 2
FILING DATE: CURNOWN>
PPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                         NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                               TOPOLOGY: linear;

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MBER: 60/092,339
1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 8, Application US/09756875 Patent No. US20020150990Al GENERAL INFORMATION:
                                                                                                                                                                                                                                     LENGTH: 74 amino acids TYPE: amino acid
                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PatentIn version 3.0
                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION NUMBER:
PRIOR FILING DATE: 1999-(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQ ID NOS:
SOFTWARE: Patentin v.
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
```

```
APPLICANT: MAERTENS, GEERT
STUTYER, LIEVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO:
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                             Gaps
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
A APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                      2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 121; DB 10;
Pred. No. 1.1e-09;
4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                     WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                      111:1:1:1111 111::1111 111111 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                               1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPKROTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
US-09-851-138-46
US-09-851-138-46
; Sequence 46, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
APPLICANT: MERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: TEXAS.
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 137 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHI
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 28 Jun 19
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                        AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AGENTS NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 22; Conserva
                    22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-851-138-46
                      Matches
                                                                                                   g
                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 113;
                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC Compatible
COMPATION SYSTEM: PC-DOS/MS-DOS
SOFTMARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
A APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 121; DB 10;
Pred. No. 8.8e-10;
4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 121; DB 10;
Pred. No. 9.2e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: Dathogenic strain of the TITLE OF INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
PRIOR APPLICATION NUMBER: EP 00402225.7
                                                                                                                   CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
PPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
NAME: KAMMEKER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 108 amino acids
                                                                                                 SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81.2%;
78.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                 NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
```

Gaps

ö

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81.2%; Score 121; DB 11; Length 166; 78.6%; Pred. No. 1.4e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
Pred. No. 1.4e-09;
4; Mismatches 2;
                                                                    1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                 1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                              Sequence 152, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Sequence 42, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
  Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
INFORMATION FOR SEQ ID NO: 4
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 22; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                           US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF INVENTION: 270
COMPUTER READABLE FORM:
MEDIOM TYPE: RIOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.2%; Score 121; DB 10; Length 138; 78.6%; Pred. No. 1.1e-09;
                                                                                                                                    ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: ISO
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 11;
                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 08/836,075
FILING DATE: <U0known>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4; Mismatches
                       WHITE & DURKEE
                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81.2%; Score 121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |||:|:||||| ||||| PKPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 138 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
                          ADDRESSEE: ARNOLD, WH
STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
    CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 78.6
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                      COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
```

Gaps

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
0
                                               ö
Query Match 81.2%; Score 121; DB 11; Length 169; Best Local Similarity 78.6%; Pred. No. 1.4e-09; Matches 22; Conservative 4; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 121; DB 11; Length 169;
Pred. No. 1.4e-09;
4; Mismatches 2; Indels (
                                                                                1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPKRQTKRNTLRRPKNVKFPAGGQIVG 28
                                                                                                                                                                                                            US-09-899-046-44
; Sequence 44, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 169 amino acids

TYPE: amino acid

TOPOLOGY: linear

WOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 81.2%;
Best Local Similarity 78.6%;
Matches 22; Conservative
                                                                                                                                                                                            RESULT 15
                                                                                     ò
                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

Search completed: August 7, 2003, 12:01:14 Job time: 14.3636 secs

5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32

g

Non-A, Non-B Hepat Peptide VIIIE base Anti-HCV antibody

Prototype peptide

HCV core-envelope HCV core-envelope HCV core-envelope

Hepatitis C virus Antigenic epitope HCV core protein N

Hepatitis C virus Protein encoded by Hepatitis C virus

DNA fragment HC360 C11-21 NANBH-speci

Hepatitis type C Hepatitis C virus

New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual

(USSH ) US DEPT HEALTH & HUMAN SERVICES

Khudyakov YE;

Fields HA,

WPI; 1999-204671/17.

CN14 fragment of H HCV capsid peptide HCV capsid peptide NCV capsid peptide NOn-A, non-B hepat Hepatitis C virus Hepatitis C virus

AAR30688 AAR30689 AAR54065 AAW37380 AAW66083

100 100 100 100 100 100

100.0 92.7 92.0 92.0 92.0 92.0 92.0

138 138 138 138 138 138

AAY06681 AAR92985 AAW06487 AAR30687

NC mosaic protein Hepatitis C virus

Description

Ωī

Length DB

Ouery Match 1

Score

Result Š

97US-0921887

25-AUG-1997;

Scoring table:

Searched:

Minimum DB Maximum DB

Database

score:

Title: Perfect ;

Sequence:

protein

Run on:

```
HCV core protein p
PT-NANB viral stru
Hepatitis C virus
Antigen pHCalOl.
Antigen pHCalOl.
HCV fragment 1 / I
HCV fragment 2 / I
Branched peptide H
                                                                                                                                                                                                                                                                                            HCV bait polypepti
Protein encoded by
Hepatitis C virus
HCV bait polypepti
HCV capsid core pr
HCV core-envelope
Hepatitis C virus
Human hepatitis C
Human hepatitis C
HCV core-envelope
HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NC mosaic protein amino acid fragment I.
                                                                                                                                                        AAR29160
AAR29160
AAR29161
AAR41422
                                                                                                                                                                                                      AAR41423
AAR51389
AAR66619
                                                                                                                                   AAW01865
AAR12597
                                                                                                                                                                                                                                                                                            ABB77253
AAB18537
                                                                                                                                                                                                                                                AAR40978
                                                                                                                                                                                                                                      AAB31695
                                                                                                                                                                                                                                                                                   AAW30597
                                                                                                                                                                                                                                                                                                                                                                                                AAR96548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY06681 standard; Protein; 28
  2200
2200
2200
2200
2200
2200
2200
2300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-JUN-1999 (first entry)
 Hepatitis C virus
 21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY06681;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY06681
ID AAY
  August 7, 2003, 11:05:37 ; Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                            3: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:
/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:
                                                                                                                                                                                                                                                                                                                                         /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.
/SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.
                                                                                                                                                                                                                           1107863
          5.1.6
Compugen Ltd.
                                                                                                                                                                                                                         al number of hits satisfying chosen parameters:
                                                                                                                                            1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                    1107863 seqs, 158726573 residues
          GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                       protein search, using sw model
                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                              A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                               seq length: 0 seq length: 2000000000
                                                                                                                        US-09-491-146A-31
```

```
Sequence
                                       Query Match
                                                                Matches
                                                                                                                                                       RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
AAR30687
     X S
                                                                                                                a
                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                 The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, sellva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                       Length 28;
                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                     100.0%; Score 150; DB 20;
100.0%; Pred. No. 2.2e-14;
ilve 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus isolate SA6 core protein.
                                                                                                                                                                                                                                                                                                      1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                      (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 4; Page 221-222; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                            AAR92985 standard; Protein; 191 AA
                                                                                                                                                                                                    comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Purcell RH;
             Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                  Local Similarity 100. es 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1996-139709/14.
N-PSDB; AAT16659.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                             28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9605315-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                              02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-FEB-1996,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; E1; er
hepatitis.
                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                      wery Match
                                                                                                                                                                                                                                                                                                                                                                                                                     AAR92985
                                                                                                                                                                                                                                                                              .4atches
                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
AAR92985
ID AAR9
                                                                                                                                                                                                                                                                                                                              q
```

```
This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of HCV infection and to raise antibodies against it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Detection of hepatitis C virus - using oligopeptide fragment of HCV
                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus; HCV; open reading frame; "common" sequence;
                                     Length 191;
                                                                                                                                                                                                                                                                                                                                                                          CN14; CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 36;
                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                   Score 139; DB 17;
Pred. No. 6.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 16;
Pred. No. 1.5e-12;
2; Mismatches 1;
                                                                     2; Mismatches
                                                                                                          28
                                                                                                                           5 PKPQRKTQRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                    Æ
                                                                                                                                                                                                                                                                                                                                           CN14 fragment of HCV core region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2
                                                                                                                                                                                                                                AAW06487 standard; Protein; 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       38
                                 92.7%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.0%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR30687 standard; peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV capsid peptide No. 23.
                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (IMMO ) IMMUNO JAPAN KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1995-047903/07.
N-PSDB; AAT45055.
                                            Local Similarity
es 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 25; Conserv
191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-MAY-1993;
                                                                                                                                                                                                                                                                                                       31-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           core region
                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic.
                                                                                                                                                                                                                                                                   AAW06487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR30687;
                                                                                                                                                                                                                                                                                                                                                                                             antibody.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seguence
```

capsid non-A,

```
Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                        Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                               The sequences given in AAR30665-89 represent fragments of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.0%; Score 138; DB 14;
89.3%; Pred. No. 1.6e-12;
tive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                              (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig 1F; 66pp; English
                                                   Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Todd JA;
                  (BAXT ) BAXTER DIAGNOSTICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR30689 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91US-0714471
91US-0718052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV capsid peptide No. 25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003 (updated)
11-MAY-1993 (first en)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Leahy DC,
                                                   Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1993-018073/02.
                                                                                      WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                  38 AA;
                                                                                                                                            virus - for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-JUN-1991;
20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Jolley ME,
                                                     Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR30689;
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                            The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoasasys for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                        Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 14; Length 38;
Pred. No. 1.6e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                (BAXT ) BAXTER DIAGNOSTICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR30688 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative 2
                                                                                                                                                                                          91US-0714471.
91US-0718052.
                                                                                                                                                           92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91US-0714471
91US-0718052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                   Jolley ME, Leahy DC,
                                                                                                                                                                                                                                                                                                                    WPI; 1993-018073/02
                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09222571-A1.
                                                                                      WO9222571-A1
                                                                                                                                                                                            13-JUN-1991;
20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-JUN-1991;
20-JUN-1991;
                                                                                                                                                         29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23-DEC-1992.
```

Sequence

ò 셤

1E X & X O O O O O O O O O O X O

AAR30688;

ö

Gaps

ó

infection

```
The present sequence represents a Hepatitis C virus (HCV) protein sequence from the disclosure of the present specification. The present specification describes a chimeric HCV peptide antigno which comprises at least 2 peptide epitope regions from the HCV polypeptide core region, 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS4 region. The antignen binds specifically with an antibody produced by a human infected by HCV. The peptide can detect a wide range of HCV infections with high sensitivity.
                                                                                                                                                                                                                                                                                                                                                                                                         New chimeric peptide antigen derived from hepatitis C virus protein - useful for detecting HCV infections
                                                                                                                                              Hepatitis C virus; HCV; chimeric; antigen; detection; core region;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 24; 30pp; Japanese.
                                                                                                                Hepatitis C virus C-1 protein 1-43.
                      AAW37380 standard; Protein; 43 AA.
                                                                                                                                                                                                                                                                                                                96JP-0024045.
                                                                                                                                                                                                                                                                                                                                                                             WPI; 1998-022248/03
                                                                                                                                                             epitope; NS3; NS4;
                                                                                                                                                                                                                                                                                                                                              TOFU ) TONEN CORP.
                                                                                                                                                                                           Hepatitis c virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        43 AA;
                                                                                                                                                                                                                                                                                 10-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    05-MAR-1998;
                                                                                                                                                                                                                      JP09278794-A
                                                                                                                                                                                                                                                                                                                09-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9839360-A1
                                                                                    11-MAR-1998
                                                                                                                                                                                                                                                      28-OCT-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11-SEP-1998
                                                     AAW37380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW66083;
                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                      Gaps
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "cormon" sequence. These peptides are contained in the capsid protein of the virus and themselves contained in the capsid protein of the virus and themselves contain epitope antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA coding a Non-A, non-B hepatitis virus antigen - useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NSI, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; core; ENV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                   Score 138; DB 14; Length 38;
Pred. No. 1.6e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Non-A, non-B hepatitis virus coreI region fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%; Score 138; DB 15;
89.3%; Pred. No. 1.6e-12;
ive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TOKR-) 2H TOKYOTO RINSHO IGAKU SOGO KENKYUSHO
                                                                                                                                           (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                 1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                 5. PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKTERNTNRRPODVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 11; 22pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SANW ) SANWA KAGAKU KENKYUSHO CO. (TOFU ) TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                        AAR54065 standard; Protein; 38
                                                                                                                                                                                                   Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   detecting HCV within serum
                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 92.0°
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1994-205028/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                         38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAQ64067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JP06141870-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                   14-FEB-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-MAY-1994.
                                                                                                                             antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                      AAR54065;
                                                                                                                                                                                                                                                                                                                                          AARS4065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        XTTXXXXXXXXXXXXXX
      ò
                                                                                                                                                                                                                                                                 ò
```

```
ö
                                 Gaps
                                                                                                                                                                                                      antigenic; nucleocapsid; p21 protein; alpha-helical; immunogen; antibody; hepatitis C virus; epitope; N-terminal fragment.
                                 ö
             Length 43
                                 Indels
                                                                                                                                                                                   Hepatitis C virus p21 protein N-terminal fragment.
           92.0%; Score 138; DB 19;
89.3%; Pred. No. 1.8e-12;
11ve 2; Mismatches 1;
                                                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                              Ź
                                                                                                                       AAW66083 standard; peptide; 44
Query Match
Best Local Similarity 89.3%,
-hes 25; Conservative
                                                                                                                                                                                                                                                                                                  98WO-FR00442
                                                                                                                                                                                                                                                                                                                     97FR-0002878
                                                                                                                                                               16-NOV-1998 (first entry)
                                                                                                                                                                                                                                                                                                                     05-MAR-1997;
```

5

ö

```
WO200031130-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        н
                                                                                                                                                                                                                                                                                                                                                                        11-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dalbon P,
                                                                                                                                                                                                                                                                                                                                              AAY94410;
                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                      AAY94410
  X & X & C & C & C & C & X & S
                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                       The invention relates to a peptide, which is recognised by antibodies against amino acids 2.45 at the N-terminus of the core (or nucleocapsid) p21 protein of hepatitis C virus (HCV), or its variants. The peptide has a tertiary structure consisting of two alpha-helical fragments, almost perpendicular to each other in space, connected by a junction peptide. Excluded are all proteins and peptides comprising, or consisting of, the N-terminal part of p21 (starting from amino acid 1 or 2). Also new are (1) monoclonal or polyclonal antibodies produced using the peptide as an immunogen and (2) complex consisting of the peptide specifically bound to some other molecule, particularly peptide or nucleotide fragments or functionalised aromatic compounds. The peptide is used (1) as immunogen for generating antibodies and (11) for detecting and quantifying either antibodies against p21 or HCV-derlyed mRNA (by
                                                                                                                                                                                                                                                                                                                                                                                prevent HCV infections. The present sequence represents the N-terminus the core p21 protein of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                           or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Monoclonal antibodies useful for detecting and/or quantifying hepatitis
                                                                                                                                                                                                                                                                                                                                                          complex formation). Antibodies are used correspondingly to detect HCV _{
m c} related antigens. The peptide and antibodies may also be used to treat
                                                                                                            nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                           New peptide from the N-terminus of hepatitis C virus p21 protein containing the immunodominant epitope - and related antibodies, tfor diagnosis, treatment and prevention of hepatitis C infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis C virus; core protein; monoclonal antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                         Penin F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Jolivet M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 19;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus Core protein amino acids 2-45.
                                        Ladaviere L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Paranhos BC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                diagnosis; infection; sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTERNINRRPQDVRFSGGGQIVG
                                                                                                                                                 Disclosure; Page 16; 37pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                        Lacoux X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY26952 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yvon S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98FR-0003087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98FR-0003087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 92.0 Best Local Similarity 89.3 Matches 25; Conservative
                                        Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Piga N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  virus core protein
            (INMR ) BIO MERIEUX
                                                                WPI; 1998-495793/42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-530397/45
                                                                                                                                                                                                                                                                                                                                                                                                                             44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FR2775690-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Jolivet RC,
                                      Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY26952;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Epitope;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 1
                                                                                                                                                                                                                                                                                            ,888888888
```

ò g XXX

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present peptide, designated $426, corresponds to residues 2 to 45 of the N-terminus of the human hepatitis C virus (HCV) core protein. It is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoractivity with all the sera of individuals or blood samples infected with HCV and which possess antibodies directed against the core protein. An amino acid may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino acids, amine groups may be acetytated, and so on. The native antigenic sequence and its antigenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against the virus.
                                                                                   nsed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polypeptides that bind to anti-hepatitis C virus antibodies, useful for diagnosing and preventing hepatitis C infections -
                                                   Peptides AAY26949-Y26955 represent peptide epitopes derived from the N-terminus of the hepatitis C virus core protein. The peptides are used to generate monoclonal antibodies or antibody fragments specific for hepatitis C virus (HCV) core protein. The monoclonal antibodies are used for early diagnosis of HCV infections, especially by two-antibody
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human hepatitis C virus core protein N-terminus, residues 2-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                     Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 44;
                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ပဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jolivet.-Reynaud
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 138; DB 21;
Pred. No. 1.8e-12;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                     20;
                                                                                                                                                                                                                                                                                               Score 138; DB 20;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                               Pred. No. 1.86
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                         Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; Page 42; 50pp; English.
Claim 3; Page 10; 19pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY94410 standard; peptide; 44
                                                                                                                                                                                                                                                                                                     92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99WO-IB01933,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0196155
                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 89...
Thes 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                           sandwich immunoassay
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-411934/35.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                              44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  44 AA;
```

C; Virus; antibody;

g ç

```
The sequences given in AAR25929-30 and AAR29534-42 represent various clones of the polypeptide core-envelope region of the Hepatitis C Virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The RNA sequences encoding these peptides were converted into cDNA using transcriptase in the presence of one of the primer sequences given in AAQ35540-46. The cDNA sequences isolated represent different alleles of the same region of the HCV gene. The entire HCV protein (see AAR29527) is useful in the development of a diagnostic method which is more accurate and effective than conventional ones, in the detection of antibodies raised against a wide range of HCVs which have been hardly detected before. The complete gene may be used in can in vitro screening system for a substance capable of specifically supressing or controlling a proteolytic processing of a precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clone; polypeptide; core-envelope; region; Hepatitis C; Virus; HCV; HC; cDNA; transcriptase; primer; allele; diagnostic; antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New hepatitis C virus gene and its encoded protein – used for diagnosing and vaccinating against hepatitis C virus infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 55;
                  Clone, polypeptide, core-envelope, region, Hepatitis HC; cDNA, transcriptase, primer; allele; diagnostic, supress; control; proteolytic processing; precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92.0%; Score 138; DB 13;
89.3%; Pred. No. 2.3e-12;
ive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                  Seki M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Honda Y, Murakami T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 5; Page 74; 305pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR29535 standard; Protein; 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV core-envelope peptide N1-2.
                                                                                                                                                                                                                                                                                                                                                                         (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                                                                                                         91JP-0172794,
91JP-0287008,
91JP-0332329,
                                                                                                                                                                                                                    92EP-0109812
                                                                                                                                                                                                                                                       91JP-0139268
                                                                                                                                                                                                                                                                                                                                    92JP-0099957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1992-417213/51.
                                                                                               Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAQ32444.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55 AA;
                                                                                                                                                                                                                11-JUN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Teranishi Y;
                                                                                                                                                                                                                                                                                                                                    20-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
26-APR-1993
                                                                                                                                      EP518313-A2
                                                                                                                                                                             16-DEC-1992
                                                                                                                                                                                                                                                                                           07-0CT-1991
16-DEC-1991
                                                                                                                                                                                                                                                                                                                                                                                                                Hayashi N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR29535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clone;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR29535
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is the 45 N-terminal amino acids of the core protein of the human hepatitis C virus (HCV). The sequence is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoreactivity with all the sera of individuals or blood samples infected with HCV and which possess antibodies directed against the core protein. An amino acid of the present sequence may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino acids, amine groups may be acetylated, and so on. The present antigenic sequence and its antigenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against
                                                                                                                                                                                                                                                                                     Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide; hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polypeptides that bind to anti-hepatitis {\tt C} virus antibodies, useful diagnosing and preventing hepatitis {\tt C} infections -
                                                                                                                                                                                                                                                  Human hepatitis C virus core protein N-terminus, residues 1-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92.0%; Score 138; DB 21; 89.3%; Pred. No. 1.9e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jolivet-Reynaud C;
28
                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 2; Page 42; 50pp; English.
                                                                                                                                AAY94409 standard; peptide; 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV core-envelope peptide N1-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR29534 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99WO-IB01933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0196155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-411934/35.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45 AA;
                                                                                                                                                                                                                                                                                                                                                                                       WO200031130-A1
                                                                                                                                                                                                              11-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
26-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                              02-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            for detect
the virus.
                                                                                                                                                                         AAY94409;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR29534;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

ö

Gaps

ö

Indels

RESULT 13 AAR29534

DXXXXX

ð

XX So

supress; control; proteolytic processing; precursor

```
1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: August 7, 2003, 11:14:09
Job time : 38.6364 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The sequences given in AAR25929-30 and AAR29534-42 represent various clones of the polypeptide core-envelope region of the Hepatitis C clones of the polypeptide core-envelope region of the Hepatitis C virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The RNA sequences encoding these peptides were converted into cDNA using transcriptase in the presence of one of the primer sequences given in AAQ32540-46. The cDNA sequences isolated represent different alleles of the same region of the HCV gene. The entire HCV protein (see AAR29527) is useful in the development of a diagnostic method which is more accurate and effective than conventional ones, in the detection of antibodies raised against a wide range of HCVs which have been hardly detected before. The complete gene may be used in an in vitro screening system for a substance capable of specifically substance controlling a proteolytic processing of a precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Clone; polypeptide; core-envelope; region; Hepatitis C; Virus; HCV; HC; cDNA; transcriptase; primer; allele; diagnostic; antibody; supress; control; proteolytic processing; precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                New hepatitis C virus gene and its encoded protein - used for diagnosing and vaccinating against hepatitis C virus infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%; Score 138; DB 13; Length 55; 89.3%; Pred. No. 2.3e-12; Live 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                        Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 33; Page 87; 305pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR29536 standard; Protein; 55 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV core-envelope peptide S1-3.
                                                                                                                                                                                                                                                                                    (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                             91JP-0172794.
91JP-0287008.
                                                                                                                                                                                                                             91JP-0332329.
                                                                                                                                        92EP-0109812
                                                                                                                                                                            91JP-0139268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (updated)
                                                                                                                                                                                                                                                                                                                                                                           WPI; 1992-417213/51.
N-PSDB; AAQ32445.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             uery Match
Jest Local Similarity
                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         55 AA;
                                                                                                                                                                                                                                                                                                                                           Teranishi Y;
                                                                                                                                        11-JUN-1992;
                                                                                                                                                                                                                               16-DEC-1991;
20-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003
26-APR-1993
                                                                                                                                                                            11-JUN-1991;
                                                                  EP518313-A2
                                                                                                    16-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR29536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
AAR29536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
```

ò g

ö

Gaps

ö

Length 55; 1; Indels

92.0%; Score 138; DB 13; 89.3%; Pred. No. 2.3e-12; vative 2; Mismatches 1;

55 AA;

```
The sequences given in AAR25929-30 and AAR29534-42 represent various clones of the polypeptide core-envelope region of the Hepatitis C clones of the polypeptide core-envelope region of the Hepatitis C virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The RNA caquences encoding these peptides were converted into CDNA using transcriptase in the presence of one of the primer sequences given in AAQ32540-46. The CDNA sequences isolated represent different in AAQ32507. Is useful in the development of a diagnostic method which is more accurate and effective than conventional ones, in the detection of antibodies raised against a wide range of HCVs which hardly detected before. The complete gene may be used in an in vitro screening system for a substance capable of specifically superssing or controlling a proteolytic processing of a precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New hepatitis C virus gene and its encoded protein – used for diagnosing and vaccinating against hepatitis C virus infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Honda Y, Murakami T, Seki M, Takahashi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polypeptide of HCV. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 33; Page 88-89; 305pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                          (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                                                                                     91JP-0172794.
91JP-0287008.
91JP-0332329.
                                                                                                                                              92EP-0109812.
                                                                                                                                                                                                                   91JP-0139268
                                                                                                                                                                                                                                                                                                                                                                     92JP-0099957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1992-417213/51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAQ32446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hayashi N, |
Teranishi Y;
                                                                                                                                          11-JUN-1992;
EP518313-A2.
                                                                     16-DEC-1992.
                                                                                                                                                                                                                   11-JUN-1991;
                                                                                                                                                                                                                                                                                        07-OCT-1991;
16-DEC-1991;
                                                                                                                                                                                                                                                                                                                                                                     20-APR-1992
```

Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Minimum DB Maximum DB

Database :

Result

```
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nicontains: core protein
C;Species: hepatitis C virus
A;Varitet; genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41355
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                 N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                      genome
genome
genome
                                                                                                                               genome
genome
genome
genome
                                                                                                                                                                                                                                                               genome
                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                          genome
                                                                                                                                                                                                                                                                                                                                                                                                                                   C virus (genotype 2, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 138; DB 2; Length 108;
Pred. No. 2.1e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 138; DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: BMBL:229458
A; Experimental source: genotype 2, N4
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S18030
GNWVCH
S4070
JC5620
JC5620
GNWVJ3
S41352
S41351
S41351
S41365
S41365
PC2219
GNWVTC
GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein - hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
876
3010
3010
3010
3011
3011
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 25; Conserv
992.0
992.0
992.0
992.0
992.0
992.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
993.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
903.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                     qenome
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepa
hepa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein
genome polyprotein
                                                                                                                         August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein - polyprotein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             denome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
               GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                    al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                     283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                         protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S12707
PC1284
JQ1925
JQ1926
                                                                                                                                                                                                                                                                                                  Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                              seq length: 0
seq length: 2000000000
                                                                                                                                                                                                     US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR_76:*
!: pir1:*
?: pir2:*
3: pir3:*
!: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4.
341348
genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
N:Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N6
A; Variety: genotype 1, N6
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41348
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
' Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
' Verence number: S41348
ccession: S41348
                                                                                                                                                                                             $41357
genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                  5 by LiPA
                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%; Score 138; DB 2; Length 108; 89.3%; Pred. No. 2.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 108;
                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                    Ç
                                                                                                                                                                                                                                                                                                                             C, Accession: $41357
R; van Doorn, L,J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A: Description: Analysis of hepatitis C virus genotypes 1
A: Reference number: $41341
Pred. No. 2.1e-12;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ... residues: 1-108 cvan>
A;Cross-references: BMBL:229460
A;Experimental source: genotype 2, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted cMAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Molecule type: genomic RNA
A;Residues: 1-108 </Wh>
A;Cross-references: EMBL: 229451
A;Experimental source: genotype 1, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Rewords: capaid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted </WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 138; DB 2;
Pred. No. 2.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 2.16
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Mismatches
                                                                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
ilarity 89.3%;
Conservative
  88.38;
  Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatit;
N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                olecule type: genomic RNA residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                            ccession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                   ô
                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
```

```
A; Variety: genotype 5, N5
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 841371
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: 841371
A; Rocession: 841371
A; Rocession: 841371
A; Residues: 1-112 < VANN
A; Residues: 1-112 < VANN
A; Residues: 1-112 < VANN
A; Residues: BMBL: 229474
A; Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: capsid protein; core protein; polyprotein
F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 1, N1
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          V.Contains: core protein
C.Species: hepatitis C virus
A:Variety: genotype 5, N4
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C. Accession: 541341

R. Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994
A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A:Reference number: 541341
A:Accession: 541341
A:Experimental source: genotypel, NID:g443850; PIDN:CAA82582.1; PID:g443851
A:Experimental source: genotypel, NI
C:Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: EMBL:229473; NID:9443908; PIDN:CAA82611.1; PID:9443909
A;Experimental source: genotype 5, N4
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lipa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ρχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Accession: S41370
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
A; Reference number: S41341
A; Accession: S41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 2;
Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.0%; Score 138; DB 2;
89.3%; Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA A; Residues: 1-114 <VAN>
```

us-09-491-146a-31.rpr

ò ద

```
Nicontains: core protein:
Cispecies: hepatitis C virus
A;Variety; genotype 1, N10
Cispecies: hepatitis C virus
A;Variety; genotype 1, N10
C;Accession: S41342
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 1, N2
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                 A; Accession: S41342
A; Molecule type: genomic RNA
A; Residues: 1-115. VANN>
A; Experimental source: genotype 1, N10
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <4MT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: EMBL:Z29447; NID:g443856; PIDN:CAA82585.1; PID:g443857
                               genome polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C'virus (genotype 1, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Accession: S41344
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genorypes 1 to
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Experimental source: genotype 1, N2 C;Superfamily: hepatitis C vitus genome polyprotein C;Reywords: capsid protein; core protein; polyprotein F;1-115/Product: core protein *status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 2;
Pred. No. 2.2e-12;
2; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPORKTKRNTNRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 89.5.
These 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S41344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N3
C;Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Accession: 341306
A) Reference houser, J.T.
Submitted to the EMBL Data Library, January 1994
A) Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A) Reference number: S41341
A) Reference number: S41341
A) Recession: S4136
A) Rolecule type: genomic RNA
A) Residues: 1-114 </P>
A) Residues: 1-114 </P>
A) Reperimental source: genotype 5, N2
C) Superfamily: hepatitis C virus genome polyprotein
C) Superfamily: hepatitis C virus genome polyprotein
C) Reproduct: core protein; polyprotein
F):1-114/Product: core protein #status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A. Reference number: 541341
A. Accession: 541369
A. Molecule type: genomic RNA
A. Residues: 1-114 < VAN>
A. Residues: 1-114 < VAN>
A. Cross-references: EMBL: Z29472; NID: g443906; PIDN: CAA82610.1; PID: g443907
A. Experimental source: genotype 5, N3
C. Superfamily: hepatitis C virus genome polyprotein
C. Superfamily: hepatitis C virus genome polyprotein
C. Keywords: capsid protein; core protein; polyprotein
F. 1-114/Product: core protein #status predicted < MAT>
                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       an Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. ....bmitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                                   Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 2; L4
Pred. No. 2.2e-12;
2; Mismatches 1;
C, Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                              Score 138; DB 2;
Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 2;
Pred. No. 2.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Vicontains: core protein

C;Species: hepatitis C virus

C;Species: penotype 5, N2

variety: genotype 5, N2

)ate: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                       2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Mismatches
                                                                                                                                                                                                                           1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ch 92.0%;
1 Similarity 89.3%;
25; Conservative
                                                                                Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Accession: $41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Accession: $41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
```

ò g

þ

Ŋ

ö

Gaps

ö

1: Indels

28

Length 115;

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, NB
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 54136
B; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Reference number: 541341
A; Reference number: 541341
                                                          Gaps
                                                       ö
                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
92.0%; Score 138; DB 2; Length 115; 89.3%; Pred. No. 2.2e-12; ive 2; Mismatches 1; Indels
                                                                                                            1 PKPORKTERNTNRRPODVRFSGGGOIVG 28
                                                                                                                                                            32
                                                                                                                                        A; Molecule type: genomic RNA
```

RESULT 10

δ g

```
A;Variety: genotype 1, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: 229450; NID: 9443862; PIDN: CAA82588.1; PID: 9443863
A; Experimental source: genotype 1, N5
A; Experimental source: genotype 1, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                               Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                           polyprotein - hepatitis C virus (genotype 1, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.0%; Score 138; DB 2;
89.3%; Pred. No. 2.2e-12;
tive 2; Mismatches 1;
            PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: August 7, 2003, 11:21:49
Job time : 10.5455 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                  C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                            A; Reference number: S41341
A; Accession: S41347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                             N;Contains: core protein
                                                                                                                                                                                                                                                          C; Accession: S41347
            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pecies: hepatitis C virus
A;Variety: genotype 1, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: $4134
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: $41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: Z29453; NID: 9443868; PIDN: CAA82591.1; PID: 9443869
A; Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: genomic RNA
A;Residues: 1-115 <VAN>
A;Cross-references: EMBL:229448; NID:9443858; PIDN:CAA82586.1; PID:9443859
A;Experimental source: genotype 1, N3
A;Experimental source: genotype 1, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <AAI>
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N:Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
                                                                                                                                                                               Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 2; Length 115;
Pred. No. 2.2e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 115;
                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Axcession: S41354
A;Molecule type: genomic RNA
A;Residues: 1-115 <VANA
A;Residues: 1-115 <VANA
A;Experimental source: genotype 2, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: core protein; core protein
F;1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 2;
Pred. No. 2.2e-12;
2; Mismatches 1
                                                                                                                                                                               Score 138; DB 2;
Pred. No. 2.2e-12;
2; Mismatches 1
                                                                                                                                                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                           1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.0%;
89.3%;
                                                                                                                                                                            Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ontains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

Gaps

ö

PKPQRKTERNTNRRPQDVRFSGGGQIVG 28 

ö

Gaps

ö

Length 115; Indels

```
GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 11:05:41; Search time 4.90909 Seconds

(without alignments)

268.226 Million cell updates/sec

Title: US-09-491-146A-31
Perfect score: 150
Sequence: 1 PKPORKTERNINRRPQDVRFSGGQIVG 28
```

Searched: 127863 seqs, 47026705 residues
al number of hits satisfying chosen parameters: 127863

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	Description	59 hepatitis c						ď	c	h genome	c		h genome	4	æ	0	• · ·	8 oncorhynchu		1 nicotiana t		_				-		7 vibrio para	σ	'n	و	'n	4	5 haemophilus
	Descr	P27959	001404	001403	P27960	P27961	P26663	P26662	P27958	P26660	P26661	P29846	900269	P26664	014738	028653	P28276	P16058	P21573	080361	P45978	P25549	097a9	8n660	Q9ez11	Swu80	P17889	087807	P0505	P1064	P4003	01341	12	P4433
CTIVE	ID	POLG_HCVJ2	POLG HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCVJT	POLG_HCV1	2A5D_HUMAN	2A5D_RABIT	IE63_HSV2H	ESR1_ONCMY	YB1_XENLA	RK4_TOBAC	SCD6_YEAST	ASLA_ECOLI	TRI_THEVO	DAPB_STAAM	DAPB_STAAU	DAPB_STAAW	IF2_BACSU	GUAA_VIBPA	CMGA_BOVIN	CMGA_HUMAN	GIP2_YEAST	AREA_ASPOR	TRP2_MOUSE	GUAA_HAEIN
	DB	-	Н	Н	-	-	٦	-		٦	-	4	-		٦	Н		-	-	-	-	~	-	Н	-	-	-	7	7	П	-	-		-
	Query Match Length	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3010	3011	602	586	512	622	303	282	349	551	1030	240	240	240	716	517	449	457	548	866	1172	523
æ	Query Match	92.0	~	92.0	N	2	92.0	92.0		92.0	ä	87.3	86.7	86.7	35.7		34.7	m	m	32.0	32.0	31.3	31.3	30.7	30.7	30.7	30.7	30.3	30.0	30.0	30.0	30.0	30.0	29.7
	Score	138	138	138	138	138	138	138	138	138	138	131	130	130	53.5	52.5	52	20	49.5	48	48	47	47	46	46	46	4	45.5	45	45	45	45	45	44.5
,	Result No.	П	7	9	*	Ω,	•	7	ω	ດ ີ	10	11	12	13	14	15	16	17	18	19	50	21	22	23	24	25	56	27	28	29	30	31	32	33

Q9cnx8 pasteurella	Q8xab5 escherichia	P04079 escherichia	Q8z4q3 salmonella	062644 bos taurus	Q8tdr4 homo sapien	P19679 spiroplasma	P26686 drosophila	P27280 epizootic h	Q02354 saccharomyc	-	_
GUAA_PASMU	GUAA_ECOL6	GUAA_ECOLI	GUAA_SALTI	LCT2_BOVIN	TCPL_HUMAN	RS2_SPICI	SR55_DROME	VNS2_EHDV2	YD49_YEAST	NCAP_CVM3	NCAP_CVMA5
П,			-	-	-	Н	~	r:H	٦	-	7
523	525	525	525	151	215	359	375	376	440	454	454
		7	7	m	m	m	m	۳	٣	m	m
29.7	7.67	29.	29	29.	29	59	29	29.	29.	29.	29
44.5 29.7							44 29		44 29.	44 29.	44 29

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-9101900; Pubmed-1383400;
MEDLINE-91019030; Pubmed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MNNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus (isolate HCV-476) (HCV).
Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
     NONSTRUCTURAL PROTEIN NS1 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein.
REMOVED FROM CAPSID PROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                             (POTENTIAL)
                                                                               (POTENTIAL)
                                                                                                                                                         (POTENTIAL)
                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ul-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein)
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein El (GP68) (GP70) (NSI) (Fragment).
Hepatitis C virus (isolato contains)
                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                              943F31E3514CDEF3 CRC64;
                                                                          N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 1;
Pred. No. 8e-13;
2; Mismatches
                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                            55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                      llarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10688; BAA01530.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
  5513
369
223
233
233
233
330
442
443
448
                                                                                                                                                                                                                                                                                     448
513
513 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
ses 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVH4
001404;
                                                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                           TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
ID AC DE PAC DE
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CAPSID PROTEIN C (POTENTIAL).
MATAIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUL-1993 (Rel. 26, Created)
01-JUL-11993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatítis C virus (isolate HCV-KF) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        First our Diagrams of Control of Procession 
                                                                                                                                                                                                                                                                                                                                                                                                       Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                        56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 1;
Pred. No. 8.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                       92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; JQ1925; JQ1925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVHK
                                                                              CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                       CARBOHYD
                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVHK
     à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   use by non-profit institutions as long as its content is in way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                           CAPSID PROTEIN (POTEWIAL).
MATRIX PROTEIN (POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTEWIAL).
BY SIMILARIY.
N-LINKED (GLCNAC. ...) (POTEWIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      There are no restrictions on ng as its content is in no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein); Matrix
E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VITOLOGY 188:331-341(1992).

-i- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDOPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEIN BY PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-92230232; PubMed=1314459; Okamamoto K., Lizuka H., Okama S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Fill-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.":
BY THE
                                                                                                                                                                                                                                                                                                                                                                       ö
PROTEIN C
                                                                                                                                                                                                                                                                                                                                    Length 520;
                                                                                                                                                                                                                                                                                                              Score 138; DB 1; Length 52. Pred; No. 8.1e-13;
                                                                                                                                                                                                                                                                                                   56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core
protein (Envelope protein M); Major envelope protein
proteins NS1 and NS2] (Fragment).
                  AMINOPEPTIDASE
FROM CAPSID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  737 AA.
                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                                    92.08;
                                                                                                                                                                                                                                                                                                                                                       89.38;
                                                                                                                                                                                                                                                                                                                                                       est Local Similarity 89.3 atches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVJ5
                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                                          CARBOHYD
                                                                                                    TRANSMEM
                                                                                                                                       CARBOHYD
                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                    luery Match
                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      Matches
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                           CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSL/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Last sequence update)
16-C77-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       proteins NS1 and NS2] (Fragment).

Hepailtis C virus (1solate HC-J7) (HCV).

Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -1. FUNCTION: THE SMALL PROPEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION. 1. SUBBUNIT: THE VIRION OF THIS VIRUS IS A NICLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-92330232; PubMed-1314459; Okamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Chanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a heatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 737;
                                                                                                 protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3AF699D82AD501B1 CRC64;
                                                                                                                                                                       CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 1; Le
Pred. No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GECNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                737 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPORKTERNINRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPORKTKRNTNRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat pr
Transmembrane; Nonstructural prote
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Virology 188:331-341(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 89.3
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --
                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              datches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein El (GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P65); Nonstructural protein 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 138; DB 1; Length 737; Pred. No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Mismatches
                                                                                                                                                                                                                                                                InterPro; IPR002523; HV_capsid.
InterPro; IPR002523; HV_capsid.
InterPro; IPR002513; HV_car.
InterPro; IPR002519; HCV_car.
InterPro; IPR002519; HCV_car.
InterPro; IPR002519; HCV_car.
InterPro; IPR002519; HCV_capsid; 1.
Ffam; PF01542; HCV_capsid; 1.
Ffam; PF01543; HV_car.
IPR001599; HCV_ar.
IPR01560; HCV_ar.
IPR01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
                                                                                                                                                                                                                                                EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 89.3
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVBK
ID POLG_HC
DF CLG_HC
DT 01-AUG
DT 01-AUG
DT 15-SEP
DE GENOME
DE (GP6)
DE (EC3)
DE (EC3)
DE (SC4)
DE NSFB (
OS HEPATI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Watches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDILINE-9701508B; PubMed-8651916,
MEDRETAL, Parge H. E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.; The crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                 MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by CAMP-dependent protein kinase.";
Eur. J. Blochem. 237:611-618(1996).
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                        from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M58335; AAA72945.1; -.
                                                                                                                                                               SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-JAN-98.
08-APR-98.
15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-APR-02.
26-JUN-00.
20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        interPro; IPR001410;
interPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10UV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1GX6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1A10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LNS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1037;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB;
```

```
1688
                                                                                                                                                                                                                                                                                                                        ¥
                                                                                                                                                                                                                                                                                                  1680
3010
                  1120
1122
1122
11139
11162
11168
1117
11187
11189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJA
P26662;
                                                                                                                                                                                                                                                                                                  STRAND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                       Query Match
STRAND
STRAND
STRAND
STRAND
TURN
                                                                                     STRAND
STRAND
HELIX
TURN
TURN
STRAND
                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                               STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | DECT | POLG | 
                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                      TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
 á
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROFEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NONSTRUCTURAL PROTEIN NA4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; 3D-structure.

1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -LINKED (
-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                           InterPro; IPR002166; HCV_RGRP,
InterPro; IPR002166; HCV_RGRP,
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; INCV_COTE; I.
Pfam; PP01560; HCV_COTE; I.
Pfam; PP01560; HCV_NS2; I.
Pfam; PP01000; HCV_NS3; I.
Pfam; PP01001; HCV_NS3; I.
Pfam; PP01001; HCV_NS4a; I.
Pfam; PP01001; HCV_NS4a; I.
Pfam; PP01001; HCV_NS4a; I.
Pfam; PP0100998; VITA_RGRP; I.
ProDom; PP0186662; HCV_NS1; I.
ProDom; PP0186662; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                         645
2041
22240
22240
22529
22529
22529
10047
10047
10063
10076
10087
10097
11097
11097
                                IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         623
645
2041
2077
2240
2529
2788
1031
                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TURN
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
DR R DD R DD R DD R DD R R DD R DD
```

```
7. 01-AUG-1992 (Rel. 23, Created)
7. 01-AUG-1992 (Rel. 23, Last sequence update)
7. 01-AUG-1992 (Rel. 23, Last sequence update)
7. 28-FEB-2003 (Rel. 41, Last annotation update)
7. 28-FEB-2003 (Rel. 41, Last annotation update)
8. 28-F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  {RNA}{N}.
SUBJUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SECUENCE FROM N.A.
MEDLINE-91086550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
Japanese patients with non-8 hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138; DB 1; Length 3010;
Pred. No. 5.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match 92.0%;
Local Similarity 89.3%;
hes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
11133
11136
11161
11161
11161
11171
11186
11188
11197
11197
```

```
Ž
                                                                                                                                                                        Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11108;
                                 623
645
2041
2077
2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of unwinding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [RNA] (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                  POLG_HCVH
P27958;
    CARBOHYD
CARBOHYD
                                                CARBOHYD
CARBOHYD
CARBOHYD
                                 CARBOHYD
                                                                                             CARBOHYD
                                                                                                              CARBOHYD
                                                                                                                            SEQUENCE
                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             virus:
                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                 POLG_HCVH
                                                                                                                                                                                                                                                                                                 RESULT
    FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                  ò
                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS institute of Bioinformatics and the EWBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CELLULAR ANTHOPRETIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001650; Helicase.C.
InterPro; IPR001650; RNA_POL_DS_PS.
InterPro; IPR001694; RNA_POL_DS_PS.
Pfam; PP01543; HCV_Cappid; 1.
Pfam; PP01542; HCV_Core; 1.
Pfam; PP01560; HCV_NS; 1.
Pfam; PP01560; HCV_NS; 1.
Pfam; PP0106; HCV_NS; 1.
Pfam; PP01006; HCV_NS; 1.
Pfam; PP01001; HCV_NS; 1.
                                                                                                                                                                                                                                                               HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                              HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002166; HCV_RdRP
                                                                                                                                                                                                                                                                                                HCV_env.
                                                                                                                                                                     EMBL; D90208; BAA14233.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3010
369
369
1100
1110
1237
1319
1319
1319
2250
2250
4413
4423
4430
532
                                                                                                                                                                                                                                                                                            InterPro; IPR002519;
InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                              InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002868;
                                                                                                                                                                                                                                                 InterPro; IPR001410;
                                                                                                                                                                                                                                                                               InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                             InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001490;
                                                                                                                                                                                      PIR; A39253; GNWVCJ
                                                                                                                                                                                                HSSP; P26663; LJXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1863
2014
2014
1083
11107
11107
12165
1316
209
209
209
4417
4417
4417
532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
 THE PROPERTY OF THE PROPERTY O
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-GP-60 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (RC 3.4.99-.); Protease/helicase NS3 (P70) (Hepacivizin)
(EC 3.4.99-.); Protease/helicase NS3 (P70) (Hepacivizin)
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Blol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-90154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonuclectide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ပ
                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
  (POTENTIAL)
                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                       MW; AA993794F46DB185 CRC64;
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
                                                                                                                                                                                                                  Score 138; DB 1;
Pred. No. 5.5e-12;
                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                           3011 AA
                                                                                                                                                                                                                                         Pred. No. 5.56
2; Mismatches
                                                                                                                                                                                                                                                                                                      1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                            92.08;
                                                                                                                                                                       327017
                                                                                                                                                                                                                                         89.3%;
                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Structure 6:89-100(1998).
                                           623
645
2041
2077
2240
2788
```

us-09-491-146a-31.rsp

```
POTENTIAL)
(POTENTIAL)
   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   772CBB29CCD94753 CRC64;
    (GLCNAC
    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   327142
                                                                                                                                         AA:
                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                           CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
TURN
HELIX
TURN
STRAND
   CARBOHYD
CARBOHYD
CARBOHYD
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE SENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROCIFIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY UJ9. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY UJ9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΒY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C CELLULAR AMINOPEPTIDASE. CAPSID PROTEIN C. ENVELOPE GLYCOPROTEIN E1. ENVELOPE GLYCOPROTEIN E2. PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5B.
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMIL DECH BOX.
                                                                                                                                                                                                                             TREANSFAC, 1001;

InterPro; IPR001410; DEAD.

InterPro; IPR00151; HCV_capsid.

InterPro; IPR00251; HCV_capsid.

InterPro; IPR00251; HCV_capsid.

InterPro; IPR002519; HCV_capsid.

InterPro; IPR002519; HCV_NS1.

InterPro; IPR00140; HCV_NS2.

InterPro; IPR00140; HCV_NS3.

InterPro; IPR00140; HCV_NS4.

InterPro; IPR00140; HCV_NS4.

InterPro; IPR00166; HCV_NS4.

InterPro; IPR00166; HCV_NS5.

InterPro; IPR001650; Hellcase_C.

InterPro; IPR001650; Hellcase_C.

InterPro; IPR00169; RNA_pol_DS_PS.

InterPro; IPR00169; RNA_pol_DS_PS.

InterPro; IPR00169; HCV_core; I.

Pfam; PP01539; HCV_core; I.

Pfam; PP01539; HCV_core; I.

Pfam; PP01506; HCV_NS3; I.

Pfam; PP01501; HCV_NS3; I.
                                                                                                                                                         EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB; 1HEI; 25-NVY-98. PDB; 1A1V; 16-FEB-99. PDB; 1A1R; 17-JUN-98. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191
383
746
809
1026
1171
1972
3011
3011
1083
1107
1165
11107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               192
384
384
747
1027
11712
11712
12421
1347
11658
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
```

```
CHAIN
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=92044440; PubMed=1658196;

MEDLINE=92044440; PubMed=1658196;

MEDLINE=92044440; PubMed=1658196;

MEDLINE=92044440; PubMed=1658196;

Medida A., Miyakawa Y., Mayuni M.;

Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carriter: comparison with reported isolates for conserved and divergent regions ";

J. Gen. Virol. 72:2697-2704(1991).

J. Gen. Virol. 72:2697-2704(1991).

HYDROPHOBIC, SUGGESTING NS2A, NS2B, NS4A AND NS4B ARE FUNCTION: THE SMALL PROPEIRS OF FOUR PRELICATION.

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C. I. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Gil in the P6

POSSITION. Cys or Thr in Pl and Ser or Ala in Pl'.

C. I. CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                           01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 4), Last annocation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein RS1 (P32) (Rec 3.4.22...); Proteass/hellcase NS3 (P70) (Hepscivirin)
(EC 3.4.22...); Proteass/hellcase NS3 (P70) (Hepscivirin)
NS4B (P27); Nonstructural protein NS5A (P45); Nonstructural protein NS5B (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatiis C virus (isolate HC-36) (HVV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ΒX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
          Score 138; DB 1; Length 3011;
Pred. No. 5.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                         3033 AA
                                                                   1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                   HCV_capsid.

HCV_core.

HCV_env.

HCV_NSI.

HCV_NS2.

HCV_NS3.

HCV_NS4a.

HCV_NS4b.
          92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D00944; BAA00792.1; ---
PIR; JQ1303; JQ1303.
HSSP; P27958; JHEI.
MEROPS; S29.001; ---
MEROPS; U39.001; ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_c
                        Best Local Similarity 89.3
Matches 25, Conservative
                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                    POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                     POLG_HCVJ6
ID POLG_HC
AC POLG60,
DT 01-AUG-
DT 28-FEB-
DE Genome
DE Genome
DE GROWE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <u>+</u>
                                                                                                                                         RESULT
                                                                                              g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICARS NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       MART; SM00487; DEXDC; 1.

Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Goat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 CELLULAR AMINOPEPTIDASE
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 1;
Pred. No. 5.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVJ8
TD POLG_HCVJ8 STANDARD, PRT; 3033 AA. AC. AC. P26661;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
InterPro; IPR001166; HCV_RdRP.
InterPro; IPR001165; HCV_RdRP.
InterPro; IPR001050; Helicase_C
InterPro; IPR001050; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSYIR.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01559; HCV_EOR; 1.
Pfam; PF01589; HCV_EOR; 1.
Pfam; PF01058; HCV_EOR; 1.
Pfam; PF01005; HCV_EOR; 1.
Pfam; PF01001; HCV_EOR; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101
2038
2038
329165 MW. -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       733
1010
1619
1866
2017
3033
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      649 64
1091 109
2038 203
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

œ

Fri Aug

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS48 (P57); Nonstructural protein NS58 (P27); Nonstructural protein NS58 (P57); Nonstructural protein NS58 (P50); Nonstructural protein Hepatitis C virus (isolate HC-38) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-92230232; PubMed-1314459; Okamoto H., Lizuka H., Tanaka T., Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; **Full-length sequence of a hepatitis C virus genome having poor pomonology to reported isolates: comparative study of four distinct genotypes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VITOLOGY 188:331-341(1992).

-I- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

HYDROPHOBIC, SUGGESTING A ROLE IN THE VIRAL RNA REPLICATION.

-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001049; HCV_NS4.
InterPro; IPR001049; HCV_NS4.
InterPro; IPR001049; HCV_NS4.
InterPro; IPR001049; RNA_POI_DS_PS.
InterPro; IPR007099; RNA_POI_PS.
InterPro; IPR007099; RNA_POI_PS.
InterPro; IPR007099; RNA_PS.
INTERPOINTER PRO; INTERPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTEIN C AND MRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002519;
InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PR002519;
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                 Hepacivirus.
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 ( POTENTIAL).
PROTEASE/HELICASE NS3 ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 ( POTENTIAL).
ROASTRUCTURAL PROTEIN NS4 ( POTENTIAL).
RNA-DIRECTED RNA POLYMERASE ( POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Encelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                           SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-92230206; Pubmed-1314449;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 1; Length 3033;
Pred. No. 5.6e-12;
2; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                           CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-APR-1993 (Rel. 25, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVTW
P29846;
                                                                                                                                                                                                                                                                                                                                  ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

SYSTEM (BY SIMILARITY) SYSTEM (BY SIMILARITY) SYSTEM (BY SIMILARITY)

(POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL).

GLCNAC GLCNAC (GLCNAC

(GLCNAC

(POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL)

GLCNAC

(GLCNAC.

```
DEFORMATION OF STANDARD: PRT; 3010 AA.

AC 000269;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
OT 01-APR-1993 (Rel. 42, Last sequence update)
DT 01-APR-1993 (Rel. 42, Last annotation update)
DE CR 34.22.); Proteane Protein NS7 (PRO) (Hepacivirin)
DE (Rel. 34.22.); Nonstructural protein NS5A (PS6); Nonstructural protein
DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
CVITURES; SSRNA posltive-strand viruses, no DNA stage: Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                           AAD267D55CDFE215
                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                     Score 131; DB 1;
Pred. No. 6e-11;
                                                                                                    (GLCNAC
                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                             ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                               2; Mismatches
                               RELAY
                  CHARGE RELAY
CHARGE RELAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 KPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                   N-LINKED
                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
N-LINKED
                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                         2 KPORKTERNTNRRPQDVRFSGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                   87.3%;
88.9%;
                                                                                                                                                                                                                                                                                                                                                          327047
                                                                                                                                                                                                                                                                                                                                                                                Query Match 87.3
Best Local Similarity 88.9
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=31642;
                                                                                                                                                                                                                                                            623
645
2041
2077
2240
2529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
    ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                        CARBOHYD
                                                                                                               CARBOHYD
                                                                                                                             CARBOHYD
                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
POLG_HCVJT
     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                    -I-FUNCTION: THE SMALL PROTEINS NS2B, NS4B AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-I-CATALYTIC ACTIVITY: HYDROLYSIS OF four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-I-CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
-I-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN MND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-I-SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPERIDAGE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
"The Talwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA."; Virology 188:102-113(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Envelope protein; Helicase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                               or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001066; HCV_RGAP.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                         EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01006; HCV_NS4a; 1.

Pfam; PF01001; HCV_NS4a; 1.

Pfam; PF01506; HCV_NS5a; 1.

Pfam; PF00291; helicase_C; 1.

Pfam; PF00999; Viral_RGRP; 1.

ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002531; HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Core protein; Coat protein;
                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                     1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                       A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                          MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
```

ö

Gaps

ö

DB 1; Length 3010; 1; Indels

POTENTIAL)

(POTENTIAL) (POTENTIAL)

(GLCNAC

(POTENTIAL

(GLCNAC (GLCNAC

Gaps

ö

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL)

POTENTIAL)

(POTENTIAL). (POTENTIAL). (POTENTIAL).

. .) (POTENTIAL)

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                           (RNA)(N).

SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                       86.7%; Score 130; DB 1; Length 3010;
85.7%; Pred. No. 8.5e-11;
1ve 2; Mismatches 2; Indels
                                                                                                                                                                MW; 94A1C77435D642BB
                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                       (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                         PRT; 3011 AA
                                                                                                                                                                                                                                                   1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
             N-LINKED
                      448
532
532
532
540
540
556
576
576
523
5241
2240
2240
22529
277
2240
22529
277
278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                  Local Similarity 85.7 nes 24; Conservative
                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
                                                                                                                                                                                                                                                                                                                         POLG_HCV1
                                                                                                        CARBOHYD
CARBOHYD
                                      CARBOHYD
                                                 CARBOHYD
                                                            CARBOHYD
                                                                       CARBOHYD
                                                                                 CARBOHYD
                                                                                              CARBOHYD
                                                                                                                               CARBOHYD
                                                                                                                                          CARBOHYD
                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                    P26664
                                                                                                                                                                                                                                                                                                RESULT 13
POLG_HCV1
                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                           8
              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                              SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN NS. (POTENTIAL).
NONSTRUCTURAL PROTEIN NS. (POTENTIAL).
NON-STRUCTURAL PROTEIN NS. (POTENTIAL).
PROTEASE/HELICASE NS. (POTENTIAL).
NONSTRUCTURAL RROTEIN NS.4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROPEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                     LICEPPO, IPRO01490; HCV_NS4a.
LICEPPO, IPRO02868; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR007095; RNA_POI_DS_PS.
InterPro; IPR007095; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_PSvir.
Pfam; PPO1543; HCV_Core; I.
Pfam; PPO1539; HCV_Core; I.
'fam; PPO1539; HCV_Core; I.
'fam; PPO1539; HCV_NS1'.
'fam; PPO1538'."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE
ATP (PO
                                                                                                                                                                                                                                                HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01506; HCV_NSAs; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                        HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                               EMBL; D11168; BAA01943.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                            PF01560; HCV_NS1; 1.
PF01538; HCV_NS2; 1.
PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1.
PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1006
11615
12862
12003
13010
11003
11003
11003
1234
1319
1319
1319
1309
1309
1309
1309
                                                                                                                                                                                                                                                                                Interpro; IPR002531; H
Interpro; IPR002518; H
Interpro; IPR004109; H
Interpro; IPR001490; H
Interpro; IPR001490; H
Interpro; IPR002868; H
                                                                                                                                                                        PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                             MEROPS; S29.001; -. MEROPS; U39.001; -. InterPro; IPR001410;
                                                                                                                                                                                                                                                  InterPro; IPR002522;
                                                                                                                                                                                                                                                           InterPro; IPR002521;
                                                                                                                                                                                                                                                                         InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1616
1863
2014
347
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
```

```
TISSUE—Colon, Eye, and Kidney.

X MEDINE—22388257; PubMed=12477932;

X Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

X Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Altschul S., Jordan H., Moore T., Max N.F., Rubin G.M., Hong L.,

Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

Altschul M.J., Waln T.B., McKernan K.J., Malek J.A., Gunaratue P.H.,

Altalon D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Altakešley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Fetal brain; (TOTOTAL 2011);
MEDLINE-96355607; PubMed-8703017;
MCCTIGHT B., Rivers A.M., Audlin S., Virshup D.M.;
MCCTIGHT B. Rivers A.M., Audlin S., Virshup D.M.;
The B56 family of protein phosphatase 2A (PP2A) regulatory subunits encodes differentiation-induced phosphoproteins that target PP2A to both nucleus and cytoplasm.";
both nucleus and cytoplasm.";
[2]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. (ISOFORM DELTA-2), AND SEQUENCE OF 501-508; 550-559; 573-580 AND 584-601 (DELTA-1).

TISSUE-Bone marrow, and Brain cortex; MEDLINE-96159032; PubMed-6856219;

Tanabe O., Nagase T., Murakami T., Nozaki H., Usui H., Nishito Y., Hayashi H., Kagamiyama H., Takeda M.; "Molecular cloning of a 74-kDa regulatory subunit (B'' or delta) of human protein phosphatase 2A."; FEBS Lett. 379:107-111(1996).
                                                                                                                                                                                                                                                                                      014738; 000494; 000696; 015171; 002 fm.
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2003 (Rel. 42, Last annotation update)
Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, PB56 delta isoform) (PP2A, B subunit, PR61 delta isoform) (PP2A, B subunit, R5 delta isoform).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-Brain cortex;
MEDLINE-97324098; PubMed-9180267;
Tanabe O., Gomez G.A., Nishito Y., Usui H., Takeda M.;
Tanabe T., Molecular heterogeneity of the cDNA encoding a 74-kba regulatory subunit (B', or delta) of human protein phosphatase 2A.";
FEBS Lett. 408:52-56(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human).
Wasaryota: Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                       ö
                                                 Length 3011;
        327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-2).
                                             Score 130; DB 1; L
Pred. No. 8.5e-11;
3; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-3)
                                                                                                                                                                                                                                                                          602 AA
                                                                                                                            1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                 SEQUENCE FROM N.A. (ISOFORM DELTA-1).
                                             Ouery Match 86.7%;
Best Local Similarity 82.1%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                          STANDARD;
        3011 AA;
                                                                                                                                                                                                                                                                       2A5D_HUMAN
        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                              PPP2R5D
                                                                                                                                                                                                                               RESULT 14
2A5D_HUMAN
                                                                                                                              ò
                                                                                                                                                                                                                                                                            CELLULAR AMINOPERIDAGE.
CARSID PROTEIN (POTENTIAL).
MATAX PROTEIN (POTENTIAL).
MAJOR ENVELOR PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BY SIMILARITY).
(BY SIMILARITY).
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ( POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINKED
                                                                                              InterPro; IPR001410; DEAD.
InterPro; IPR001521; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002511; HCV_core.
InterPro; IPR002511; HCV_env.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR001618; HCV_NS2.
InterPro; IPR00146; HCV_NS4.
InterPro; IPR00146; HCV_NS4.
InterPro; IPR00146; HCV_NS5a.
InterPro; IPR00160; HCV_NS6a.
InterPro; IPR00160; Hellcase_C.
InterPro; IPR00160; Hellcase_C.
InterPro; IPR00169; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01542; HCV_core; Pfam; PF01542; HCV_core; 1 Pfam; PF01559; HCV_core; 1 Pfam; PF01569; HCV_NS2; 1 Pfam; PF01200; HCV_NS3; 1 Pfam; PF01000; HCV_NS4s; 1 Pfam; PF01000; HCV_NS4s; 1 Pfam; PF01000; HCV_NS5a; 1 Pfam; PF00190; Viral_RGRP; 1 ProDom; PD180062; Viral_RGRP; 1 SMART; SM00487; DEXDC; 1
                                                                                                                                                                                                                                                                                                                                                                                         PF01543; HCV_capsid; 1
PF01542; HCV_core; 1.
PF01539; HCV_env; 1.
PF01560; HCV_NSI; 1.
PF01589; HCV_NS2; 1.
PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
MEROPS; S29. 001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
```

)

13

Page

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TAS.
                                                                                                                                                            SUBBURITE PRA CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 KDB CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 KDB CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES WITH A VARLETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS BY THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS AND CELL SIGNALING MOLECULES.

SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR IN INTERPHASE, ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
Missing (in isoform Delta-3).
                                                                        Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBGELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISOLG-014738-3; Sequence-VSP_005110; TISOLG-014738-1; Sequence-VSP_005110; TISSUE SPECIFICITY: ISOFORM DELTA-2 IS WIDELY EXPRESSED. ISOFORM DELTA-1 IS HIGHLY EXPRESSED IN MEARIN. INUCTION: BY RETINICA ACID; IN WENROBLASTOWA CELL LINES. PTM: AT LEAST ISOFORM DELTA-1 IS PHOSPHORYLATED ON SERINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8 X 2 AA APPROXIMATE TANDEM REPEATS
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MIM; bulo46; ...
GO; GO:0005634; F:protein phosphatase type 2A, regulator acti..
GO; GO:0008601; F:protein phosphatase type 2A, regulator acti..
GO; GO:0007399; P:neurogenes1s; TAS.
InterPro; IPR002554; B56.
Pfam; PF01603; B56; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Nuclear protein; Phosphorylation; Alternative splicing; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /FTId-VSP_005110.
Missing (in isoform Delta-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FISF71AF4E565387 CRC64;
                                                                                                                                                                                                                                                                                                                                                                     Event-Alternative splicing; Named isoforms-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                               IsoId-Q14738-2; Sequence-VSP_005111;
                                                                                                                                                                                                                                                                                                                                                                                                           IsoId-Q14738-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AB000634; BAA20381.1; -. AB000635; BAA20382.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   D78360; BAA11372.1; -. BC001095; AAH01095.1; -. BC001175; AAH01175.1; -. BC010692; AAH10692.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69991 MM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; L76702; AAB69751.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; S68686; S68686.
Genew; HGNC:9312; PPP2R5D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      530
565
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Name-Delta-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                Name-Delta-2;
                                                                                                                                                                                                                                                                                                                                                                                            Name-Delta-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      523
548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85
                                                                                                                                                   COMPARTMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESIDUES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MIM; 601646;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3K; Q14738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; EMBL; EMBL; EMBL; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
```

DB 1; Length 602;

Score 53.5; DE Pred. No. 3.1;

35.7%; 31.6%;

Best Local Similarity

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
       ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Multigene family; Nuclear protein; Repeat.
DOMAIN 21 36 8 X 2 AA APPROXIMATE TANDEM REPEATS OF Q-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPOSED OF A 56 KDB CAPALYTIC SUBBNIT (SUBBNIT C) AND A 65 KDB COMPOSED OF A 56 KDB CAPALYTIC SUBBNIT (SUBBNIT C) AND A 65 KDB CONSTANT REGULATORY SUBUNIT (PR65 or SUBBNIT A), THAT ASSOCIATES WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATES WITH THE CORE DIENE INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS B (THE R2/B/PR55/B55, R3/B*, //PR72/PR130/PR59 AND R5/B*/B56 FAMILIES), THE 48 KDB VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS, AND CELL SIGNALING MOLECULES.

-1- SUBCELLULAR LOCATION: NUCLEBAL (By similarity).
-1- TISSUE SPECIFICITY: Highly expressed in brain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-96161994; PubMed-8576224;
CSortos C., Zolnierowicz S., Bako E., Durbin S.D., Depaoli-Roach A.A.;
"High complexity in the expression of the B' subunit of protein
phosphatase 2A0. Evidence for the existence of at least seven novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Soforms.";
J. BLOL. Chem. 271:2578-2588(1996).
J. BLOL. Chem. 271:2578-2588(1996).
-I- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
   Gaps
                                                                                                                                                                                                                                                                                                       DASD_RABIT STANDARD; PRT; 586 AA.

028653; 028655;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
5crine/threonine protein phosphatase 2A, 56 KDa regulatory subunit,
5crine/threonine protein phosphatase 2A, 56 KDa regulatory subunit,
BSG delta isoform (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, BSG delta isoform) (PP2A, B subunit, RS delta isoform) (PP2A, B subunit, RS delta isoform) (PP2A, B subunit, B'-gamma).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- CAUTION: NOMENCLATURE USED IN REF.1 REFERS TO PP2A B SUBUNIT B' GAMMA ISOFORM, WHICH IS CITED AS PP2A B SUBUNIT DELTA-PR61 ISOFORM IN LATER PUBLICATIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Oryccolagus cuniculus (Rabbit).
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
       11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 586;
       4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-New Zealand; TISSUE-Brain, and Skeletal muscle;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E149A309CDDA7435 CRC64;
                                                                       27
                                                                                                                                       81
                                                                                                                                   POPOAQSOPPSSNKRPSNSTPPPTQLSKIKYSGGPQIV
                                                                       PKPQRKTE-RNTNRRPQD------VRFSGGGQIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35.0%; Score 52.5; Di
31.6%; Pred. No. 4.3;
   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68090 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; U38193; AAC48532.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U38195; AAC48534.1; -. Interpro; IPR002554; B56. Pfam; PF01603; B56; 1.
   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  507
532
586 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-9986;
   12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            family
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PPP2R5D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                              ZA5D_RABIT
   Matches
                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                              SOT THE TENT AND DEPTH OF THE TENT AND THE T
                                                                                                                                g
                                                                       à
```

12; Conservative 11; Mismatches 4; Indels 11; Gaps Matches

ò

Q

Search completed: August 7, 2003, 11:20:04 Job time : 5.90909 secs

us-09-491-146a-31.rspt

```
hepatitis
hepatitis
                                                                                                                                                                                             hepatitis
hepatitis
                                                                                                                                                                                                                                                         hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                    hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                   hepatitis
hepatitis
                                                                                  hepatitis
                                                                                                                               hepatitis
                                                                                                                                                                hepatitis
                                                                                                                                                                             hepatitis
                                                                                                                                                                                                                                                                                       hepatitis
                                                                                                                                                                                                                                                                                                                                     hepatitis
                                                                                                                                                 hepatitis
                                                                                                                                                                                                                         hepatitis
                                                                                                                                                                                                                                            hepatitis
                                                                                                                                                                                                                                                                                                         hepatitis
                                                                                                                                                                                                                                                                                                                        hepatitis
                                                                                                                                                                                                                                                                                                                                                       hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                    hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tokita H., Kaufmann G.R., Matsubayashi M., Okuda I., Tanaka T.,
Harada H., Mukaide M., Suzuki K., Cooper D.A.;
"Hepartisis C Virus Core Mutations Reduce the Sensitivity of a
"Hepartisis C Virus Core Mutations Reduce the Sensitivity of a
Fluorescence Enzyme Immunoassay.",
J. Clin. Microbiol. 38:3450-3452(2000).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY.
IIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AB039875; BAB12423.1; -
InterPro; IPR002522; HCV_capsid,
Pfan. PF01443; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
 066863
066876
068707
068707
081780
081780
081780
081780
081780
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
081787
                                                                                                                                                                                                                                                                                                                                   084p63
084p61
094t52
094t54
084p59
094t60
084p58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11644 MW; C2BA6559BDDDFCBC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94.0%; Score 141; DB 12;
92.9%; Pred. No. 2.6e-13;
11ve 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              103 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            090756
080896
080865
080865
080860
080857
090758
080863
                                                                                                                                          080P70
080P54
090T57
080P52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
 Q68683
Q68706
Q68684
Q68707
Q68685
Q8JYSQ
Q8JYSJ
Q81264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-20440673; PubMed-10970401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2001 (TrEMBLrel. 16,
01-MAR-2001 (TrEMBLrel. 16,
01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HC-S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09E977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09E977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 1
Q9E977

      1
      1
      1
      1
      2
      2
      3
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4
      4

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SO STATE OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hepatitis
                                                                                                                          (without alignments)
281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                           August 7, 2003, 11:05:41; Search time 25.6364 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09e977
             5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                   tal number of hits satisfying chosen parameters:
                                                                                                                                                                                        150
1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                     830525 seqs, 258052604 residues
              GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                 summaries
                                                                           protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9E977
Q81481
Q8V847
Q68137
Q68311
Q68309
Q8JYR9
Q8JYRS
Q8JYRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q68712
Q68682
Q68708
Q68679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28JYR6
                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                  length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_archeap:*
                                                                                                                                                                          US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               445
61
60
74
74
74
74
74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sed
                                                                                                                                                                                          Perfect score:
                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1141
1138
1138
1138
1138
1138
1138
1138
                                                                               protein
                                                                                                                                                                                                                                                                                                                                                  2000
                                                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                                                                                                                                                                                                                                                                  Minimum Maximum
                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .
9
```

ö

K

Length 191;

191 20902 MW; 43A1B0CEFB44A0F9 CRC64;

g

```
Polyprotein.
NON_TER 191 1
SEQUENCE 191 AA;
                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                     068137
                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                     RESULT
Q68137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
068311
     FT
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Okamoto H., Crovatto M., Pozzato G., Feray C., Brechot C., Mishiro S., "Molecular and clinical characteristics of the hepatitis C virus genotype '2c' found in Italians in Italy and France."; Int. Hepatol. Commun. 3:161-165(1995).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: D31986: BAA06754-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AB061932; BAB83271.1; -
InterPro; IPR00252; HCV_capsid.
InterPro; IPR004531; HCV_capsid.
Pfam; PF01543; HCV_capsid.
Pfam; PF01542; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kato N.;
"Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 1
128 128
128 AA; 14591 MW; OF5F0083F8ED0B5A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08V847;
U-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94.0%; Score 141; DB 12; 92.9%; Pred. No. 3.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
PKPQRKTKRNTNRRPQDVRFPGGGQIVG 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 PKPQRKTKRNTNRRPQDVRFPGGGQIVG 30
                                                                                                                                                                                                                                                                    Core (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26; Conservative
                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
NON_TER
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-GIUS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              lesions.
                                                                                                                                Q81481
Q81481;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8V847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                           RESULT 2 081481 1 10 081481 AC 08148 AC 0814 A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
Q8V847
ID Q8V8
```

a

STT WE DRA DECCO

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-93376778; Pubmed-8396266;

A MEDLINE-93376778; Pubmed-8396266;

Bukh J., Purcell R.H., Miller R.H.;

Bukh J., Purcell R.H., Miller R.H.;

At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative El gene of isolates collected worldwide.";

Proc. Natl. Acad. Sci. U.S.A. 90:8224-8238(1993).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPOPOTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN AND GIVCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).

CENT. ORD. STATE OF COMPLEX OF PROTEIN C. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).

RICEPPO: IPRO0252; HCV_capsid.

RICEPPO: IPRO0252; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-SA6;
MEDDINE-92279243; PubMed-1317578;
Bukh J., Purcell R.H., Miller R.H.;
Sequence analysis of the 5' noncoding region of hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92.7%; Score 139; DB 12: Length 191; 89.3%; Pred. No. 1e-12; 11ve 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-94336721; PubMed-8058787; Bukh J., Purcell R.H., Miller R.H.; Sequence analysis of the core gene of 14 hepatitis C virus genotypes.";
                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20682 MW; OEB8000CF1F902EB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
94.0%; Score 141; DB 12; 92.9%; Pred. No. 5e-13;
                                                                                                                                                                                                                                                                                                                                                   PRT; 191 AA
                                                             1; Mismatches
                                                                                                                                                               5 PKPQRKTKRNTNRRPQDVRFPGGGQIVG 32
                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPORKTORNTNRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.3'
Matches 25; Conservative
                                                             26; Conservative
                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein
NON_TER 191 1
SEQUENCE 191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
```

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Shustov A.V., Gavillova I.V., Netesov S.V.;
Gavillova Depatitis C virus in Western Siberia.";
Genetic variability of hepstitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MNNA (BY SIMILARITY).
                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                             Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                            Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    B5691CBE8F26F2F1 CRC64;
                                                                                                                                          (TrEMBLrel. 01, Last sequence update) (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 1;
Pred. No. 3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 1:
Pred. No. 4e-13;
                                                                                                                                                                            Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 AA.
                                                                                          Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                         46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28
PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60
6686 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46 46
46 AA; 5129 MW;
                                                                                                                         01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-0CT-2002 (TrEMBLrel. 22, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF506616; AAM33392.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 89.30,
Best Local Similarity 69.30,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              U23747; AAA65054.1;
                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60
AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                             STRAIN-HCV-BB10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-RIG306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein.
NON_TER 46
SEQUENCE 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER
                                                                                                                                            01-NOV-1996
01-OCT-2002
                                                                                                                                                                                                                                   Hepacivirus
'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                  virus.'
                                                                                         068309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8JYR9
                                                     RESULT 7
Q68309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OBJYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GITCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN. C AND MRNA (BY SIMILARITY).

EMBL, U23744; AAA65051.1;
InterPro; IPR00252; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                      Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
EMBL; U23749; AAA65056.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=HCV-BB5;
Songsivila1 S., Kanistanon D., Kunkitti R.;
*Identification and characterisation of Thal isolates of hepatitis
                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                 Songsivilal S., Kanistanon D., Kunkitti R.; *Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.0%; Score 138; DB 12; Length 45; 89.3%; Pred. No. 2.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 138; DB 12; Length 45;
Pred. No. 2.9e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CC527167096AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45 AA; 5030 MW; CC4C21ED236AAA81 CRC64;
                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOV-1996 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TIEMBLrel. 01, Created)
01-NOV-1996 (TIEMBLrel. 01, Last sequence upda
01-OCT-2002 (TIEMBLrel. 22, Last annotation up
01-OCT-protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002522; HCV_caps1d.
Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             45 45
45 AA; 5015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best_Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                 SEQUENCE FROM N.A. STRAIN-HCV-BB37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45
                                                                                                                                                                                                                                                                                                                                                                                                                                                           45
                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                            Hepacivirus
```

SEQUENCE

SO FT

virus.

Query Match

ö 셤 068306

AC DE DE OC OC OC OC

RESULT 6 Q68306

ö

Gaps

ö

Ø

NON\_TER SEQUENCE Query Match

ST WE DRANGE COC COC REPORT OF THE PROPERTY OF

Matches

ô

ø

```
STRAIN-RIG288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-RIG289;
                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  25;
                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                Q8JYR7
                                                                                                                                             Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QBJYR6
                                                                                                            RESULT 11
Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
OBJYR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                            g
                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                 ö
   Gaps
                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                      Shustov A.V., Gavrilova I.V., Netesov S.V.;
Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCORROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
'-- SUBMUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA (BY SIMILARITY).

EMBL; AF506620; AAM33396.1;
InterPro; IPR002522; HCV_capsid.
Pfan; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                 ö
  ö
                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 12; Length 61; Pred. No. 4.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                           61
6897 MW; EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EC656DC79E8F26F2 CRC64;
                                                                                                                             01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
  1,
                                                                                                61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ž
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG
                                                                                                                     Created)
                                                                                                PRT;
                                                                                                                                                                                                                                                                                                              EMBL, AF506617, AAM33393.1, -...
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                  Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                         92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
6897 MW;
                                                                                                                   (TrEMBLrel. 22,
25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                    61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AA;
                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                            STRAIN-RIG286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                               Polyprotein.
                                                                                                                                                               Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QBJYR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
Q8JYR5
TO ORJYR
Matches
                                                                          RESULT 9
                                                                                   Q8JYR8
                    õ
                                        셤
                                                                                                         OC OC DE DIA
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN WENTELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIXCOPPOTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AFSOGGER, AMANA (BY SIMILARITY).
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid. 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROFIEN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                 Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 61;
                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EC657F706E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
          Score 138; DB 12;
Pred. No. 4.1e-13;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 12;
Pred. No. 4.1e-13;
                                                                                                                                                                                                                                                                                   61 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Mismatches
                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                   28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKTERNTNRRPQDVRFSGGGGIVG
                                                                                                                                                                                                                                                                                                                         01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last seq
01-MAR-2003 (TrEMBLrel. 23, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AFS06619; AAM33395.1; ...
InterPro; IPR002522; HCV_caps1d.
Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein (Fragment)
Hepatitis C virus.
          92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AA; 6869 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-0CT-2002 (TrEMBLrel. 22, 01-0CT-2002 (TrEMBLrel. 22, 01-MAR-2003 (TrEMBLrel. 23,
Ouery Match
Best Local Similarity 89.37
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
```

us-09-491-146a-31.rspt

```
74
                                                                                                                                                                     Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
SEQUENCE
                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                   Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                           902890
                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
Q68708
                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                               DDT THE STAND TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
STREET
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. MEDLINE-96118171; PubMed-8578855; Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P., Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus genotyping by means of 5'-UR/core line probe assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P., Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus type lb.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 138; DB 12; Length 74; Pred. No. 5e-13;
                                                                         Length 62;
                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74 AA; 8470 MW; 2AE2CF5ED0BD10F3 CRC64;
                                   EEEC656DC79E8F26 CRC64;
                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment). Hepatitis C virus type 1b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome Polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                 Score 138; DB 12;
Pred. No. 4.2e-13;
2; Mismatches 1:
                                                                                                                                                                                                                                                                                     74 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   74 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                 1 PKPORKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                     (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDLINE-96118171; Pubmed-8578855;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
89.3%;
                 62
7053 MW;
                                                                     / Match 92.0%;
Local Similarity 89.3%;
Nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                 62 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=31647;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=31647;
                 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Maertens G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
Polyprotein.
                                                                                                                                                                                                                                                                                                                         01-NOV-1996
01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
SEQUENCE
                                                                         Query Match
                                                                                                                                                                                                                                                                                 068712
Q68712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   068682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                 068712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
S T S
                                                                                                                                                   ô
                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                         22555
```

```
ö
                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDILINE-96118171; Pubmed-8578855;
Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Puig P.,
Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Maertens G.; "Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples."; Virus Res. 38:137-157(1995).
                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                  K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus type lb.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
and molecular analysis of untypeable samples.";
Virus Res. 38:137-157(1995).
-- SDBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; 138348; AAC42168.1; -- InterPro; IPRO02522; HCV_capsid.
PFam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 138; DB 12; Length 74;
Pred. No. 5e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                         Score 138; DB 12; Length 74; Pred. No. 5e-13;
                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  74 AA; 8412 MW; 2AE2DB0F976D10F3 CRC64;
                                                                                                                                                                                                                                                 74 AA; 8440 MW; 2AE2DB0F95BD10F3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           74 AA
                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: August 7, 2003, 11:19:02 Job time: 25.6364 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                         92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 92.0°
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                  1 Similarity 89.3 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          I_TaxID-31647;
```

```
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 31, Appl
Sequence 204, App
Sequence 6, Appl
Sequence 36, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 9, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 3, Appl
Sequence 26, Appl
Sequence 23, Appl
Sequence 3, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 14, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 6, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6, Application App
                                                                                                                                                                          August 7, 2003, 11:07:41; Search time 10:5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                         150
1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                        OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                            US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               %
Query
Match 1
                                                                                                                                                                                                                                                                                                         Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                                                                                                       Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
No.
```

Sequence 2, Appli Sequence 15, Appli Sequence 16, Appli Sequence 155, Appli Sequence 156, Appli Sequence 156, Appli Sequence 157, Appli Sequence 167, Appli Sequence 169, Appli Sequence 161, Appli Sequence 161, Appli Sequence 161, Appli Sequence 162, Appli Sequence 165, Appli Sequence 165, Appli Sequence 166, Appli Sequence 169, Appli	METHOD FOR MAKING THE SAM	Length 28; Indels 0; Gaps 0;
28 138 92.0 154 3 US-08-854-531-2 29 138 92.0 154 5 PCT-US95-13552-2 31 138 92.0 190 1 US-07-681-701-16 32 138 92.0 191 2 US-08-290-665A-155 34 138 92.0 191 2 US-08-290-665A-155 35 138 92.0 191 2 US-08-290-665A-155 36 138 92.0 191 2 US-08-290-665A-155 37 138 92.0 191 2 US-08-290-665A-159 38 138 92.0 191 2 US-08-290-665A-161 39 138 92.0 191 2 US-08-290-665A-161 39 138 92.0 191 2 US-08-290-665A-161 41 138 92.0 191 2 US-08-290-665A-163 42 138 92.0 191 2 US-08-290-665A-164 43 138 92.0 191 2 US-08-290-665A-165 43 138 92.0 191 2 US-08-290-665A-165 44 138 92.0 191 2 US-08-290-665A-165 45 138 92.0 191 2 US-08-290-665A-165 46 138 92.0 191 2 US-08-290-665A-165 47 138 92.0 191 2 US-08-290-665A-165 48 138 92.0 191 2 US-08-290-665A-165 49 138 92.0 191 2 US-08-290-665A-165	ALIGNMENTS  1. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 887-31  2. 881-31  2. 881-31  2. 887-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  2. 881-31  3. 881-31  3. 881-38  4. 881-38	Query March 100.0%; Score 150; DB 3; Len-Best Local Similarity 100.0%; Pred. No. 5e-15; Matches 28; Conservative 0; Mismatches 0; In

```
Sequence 6, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHEFIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
TITLE OF INVENTION: FOR DETECTING THE LATTER
  SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 89.3%; Pred. No. 1.5e-12; Matches 25; Conservative 2; Mismatchee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: WORDPERFECT 5.11
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
NAME: RICHARD W. BORK
                                                                                                                                                                                         COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                  SSEE: MORGAN & FINNEGAN F: 345 PARK AVENUE NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE/DOCKET NUMBER: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: 421792,
INFORMATION FOR SEQ ID NO: 20.
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA6
                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & F
                                                                                                                                                 ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
ZIP: 22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCT-US95-10398-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM:
                                                                                              CITY: NEI
STATE: NI
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 139; DB 2; Length 191;
Pred. No. 1.5e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R.H.
NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUC
TITLE OF INVENTION: AMINO ACID SEQUENCES;
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                          APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: MUCLEOTIDE AND DED TITLE OF INVENTION: AMINO ACID SEQUENCE TITLE OF INVENTION: CORE GENES OF ISOLUTILE OF INVENTION: AND THE USE OF REA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2026-4116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 204, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                     Sequence 204, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 2026
REFERENCE/DOCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 756-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 204:
                                                                                                                                                                                                                                                                                                                                                                                   SSEE: MORGAN & FINNEGAN 345 PARK AVENUE NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match 92.7%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA6
US-08-290-665A-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                              NEW YORK
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PCT-US95-10398-204
                                                                                                               RESULT 2
US-08-290-665A-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                      g
                    ç
```

Gaps

```
APPLICANT: DALBON, PASCA1
APPLICANT: JOLIVET, MIChAI
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALL)
TITLE OF INVENTION: FOR DETECTING THE LATTER
CORRESPONDENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                     Score 138; DB 4; Length 43;
Pred. No. 4.1e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                          Query Match 92.0%; Score 138; DE Best Local Similarity 89.3%; Pred. No. 4.1e Matches 25; Conservative 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
CLASIPITCATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZIP: 22320
COMPUTER REALABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 05346
TELECOMMUNICATION INFORMATION:
TELEPAX: (202) 672-5300
TELEPAX: (202) 672-5300
TELEFAX: (302) 672-5300

TELEFAX: 43 and oacids

LENGTH: 43 and oacids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08380160 Patent No. 6235284 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BERRIDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: OLIFF & BERI
STREET: P.O. BOX 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 44 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEX:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: peptide HYPOTHETICAL: NO ANTI-CENT
                                                                                                                                                                                                                                                                                                                          ; MOLECULE TYPE: peptide US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                            STRANDEDNESS:
                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET: 3000 K Street, N.W.
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 2000-5109
COMPUTER REDABLE FORM:
WEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OFFRATING SYSTEM: PC-DOS/MS-DOS
SOFFWARE: PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
           OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 92.0%; Score 138; DB 3;
Best Local Similarity 89.3%; Pred. No. 3.2e-13;
Matches 25; Conservative 2; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIWAKUMA, TOMIKO
APPLICANT: CHIBA, YUKie
APPLICANT: YAGI, Shintaro
APPLICANT: YAGI, Shintaro
TILLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                  CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: 30,024
TELECHONE: (703)836-6400
TELEPAX: (703)836-5787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Human Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 09-FEB-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-024045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 36, Application US/09020846 Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-terminal
                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acids
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FRAGMENT TYPE:
                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: JOLIVET, MICHEL
APPLICANT: DALLOW:
APPLICANT: DENIN, FRANCOIS
APPLICANT: DALLOW. PEANCOIS
APPLICANT: LADAVIERE, LAURENT
TITLE OF INVENTION: OOMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
TITLE OF INVENTION: TREATING AN HCV INFECTION
CURRENT FILING DATE: 1999-09-07
ERRIER APPLICATION NUMBER: PCT/FR98/00442
ERRIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATENTIN VOF: 2.1
                                                                                                                                                                                                                                                                  ö
                                                                                                                              /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the H Strain
                                                                                                                                                                                                                        Score 138; DB 3; Length 44;
Pred. No. 4.2e-13;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 4;
Pred. No. 4.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of Patent No. 6576240
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                 1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Application US/09389756 Patent No. 6576240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1, Application US/08380160 Patent No. 6235284 GENERAL INFORMATION:
                                                                                                                                                                                                                          92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: DALBON, Pascal APPLICANT: JOLIVET, Michel
FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: Hepatitis C virus
                                                                                                                                                                                                            Ouery Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 89.3°
Watches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RELEVANT RESIDUES: 2 TO 45
                                                                                      NAME/KEY: Peptide
LOCATION: 1.44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAGES: 3392-3396
                                                                                                                                                                                                                                                                                                                                                                                              JLT 7
.09-389-756-1
                                                                                                                                                                                     US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 44
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 1
                                                                        FEATURE
                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
US-07-946-054-9
Sequence 9, Application US/07946054
Sequence 9, Application US/07946054
Sequence 9, Application US/07946054
Sequence No. 5582968
Sequence 9, Application
Sequence No. 5582968
Sequence 9, Application
APPLICANT: Wang, Chang Y1
APPLICANT: Wang, Chang Y1
TILE OF INVENTION: No. 5582968e1 Branched Hybrid and Cluster
TITLE OF INVENTION: No. 5582968-B Hepatitis
TITLE OF INVENTION: No. 5582968-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
  SYNTHETIC POLYPEPTIDES BELONGING TO THE HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY FOR DETECTING THE LATTER 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 92.0%; Score 138; DB 3; Best Local Similarity 89.3%; Pred. No. 4.3e-13; Matches 25; Conservative 2; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-YAX-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
TITLE OF INVENTION: SYNTHETIC PORTITIES OF INVENTION: HERATITIS CALLE OF INVENTION: FOR DETECTION WIMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIGGE STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                          COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: Peptide
LOCATION: 1..45
                                                                                                                                                               Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN: H77
                                                                                                                                                                                                                                    22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ANTI-SENSE:
                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
```

```
Gaps
                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 61;
                                                                                                                                                                                                         DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/530,550
                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 3, Application US/08530550
; Patent No. 573631
; GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Detection of Hepatitis c inventions of Petection of Hepatitis c inventions of Sequences: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Liaa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 1,
NO. 6e-13;
                                                                                                                                                                                                     Query Match 92.0%; Score 138; DB 1, Best Local Similarity 89.3%; Pred. No. 6e-13; Matches 25; Conservative 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 92.0%; Score 138; DE Best Local Similarity 89.3%; Pred. No. 6e-1 Matches 25; Conservative 2; Mismatches
                                                                                                                                                                                                                                                                                                              1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 26, Application US/08262037; Patent No. 5747239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                34,045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
                   TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 23:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE: peptide US-08-530-550-3
                                                                                                  TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 23, Application US/08083947

Patent No. 5639594

GENERAL INFORMATION:
APPLICANT: Mang, Chang Yi
APPLICANT: Hosein, Barbara
TITLE OF INVENTION: No. 5639594el Linear and Branched Peptides Effective
TITLE OF INVENTION: in Diagnosing and Detecting No. 5639594-B Hepat
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIDLE
COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PAtentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
                                                                                                                                                                                            OURTHAND OF STSTEM: PC-DUS/MS-LUSS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/O7/946,054
FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
ATTONENTY/AGENT INPOMMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION OF 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 aming acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 1,
Pred. No. 6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
CITY: Hauppauge
STATE: New York
                                                                                                                              COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Policy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER: 2000Z
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 92.0
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                    ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: US
ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
US-08-083-947-23
                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              င်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: TBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURREWN APPLICATION DATA:
FILING DATE:
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 23, Application PC/TUS9407088
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Peptides Effective In
TITLE OF INVENTION: Diagnosing And Detecting
TITLE OF INVENTION: Diagnosing And Detecting 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.0%; Score 138; DB 5;
89.3%; Pred. No. 6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: PCT/US94/07088
FILING DATE: 22-JUNE-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC COMPALIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDERECT 5.1
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,947
FILING DATE: 28-JUNE-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELECOMMUNICATION INFORMATION
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
REGISTRATION NUMBER: 34,6
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIN, MARIA C.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                SYNTHETIC PEPTIDES SPECIFIC FOR
THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
INFECTION AND PREVENTION THEREOF AS VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 9, Application PC/TUS9308638
GENERAL INFORMATION:
APPLICANT: United Biomedical Inc.
TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
TITLE OF INVENTION: Non-B Hepatitis
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.0%; Score 138; DB 1;
89.3%; Pred. No. 6e-13;
tive 2; Mismatches
                    APPLICANT: Chang Y1 Wang and Barbara Hosein TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECITIVE OF INVENTION: THE DETECTION OF ANTIBOC TITLE OF INVENTION: INFECTION AND PREVENTION NUMBER OF SEQUENCES: 136
CORRESPONDENCES: 136
STREET: 345 PARK AVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 PKPORKTKRNINRRPODVKFPGGGOIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: WORDPERFECT 5.1 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/667,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-ADF11-1990
ATTONNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  very Match
est Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM:
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                          NEW YORK
                                                                                                                                                                                                                                                                                           CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; TOPOLOGY:
US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

REFERENCE/DOCKET NUMBER: 11
TELECOMMUNICATION INFORMATION
TELEPHONE: (212)758-4800

ADDRESSEE: UNITED BIOMEDICAL INC. STREET: 2S Davids Drive CITY: Hauppauge STATE: New York

REGISTRATION NUMBER:

INFORMATION FOR SEC ID NO:

ö

Gaps

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Hosein, Barbara
APPLICANT: Hosein, Barbara
APPLICANT: Hosein, Chang YI
TITLE OF INVENTION: Peptides Effective for
TITLE OF INVENTION: Diagnosis and Detection of Hepatitis C Infection
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P
STREET: 345 Park Avenue
COUNTRY: USA
COUNTRY: USA
COUNTRY: USA
COUNTRY: USA
COMPUTER: FLOPPY disk
COMPUTER: Tan PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
COMPUTER: BATELOATION DATA:
APPLICATION NUMBER: DCT/US95/13660
FILING DATE: 23 October 1995
CLASSIFCATION:
APPLICATION NUMBER: 08/333,573
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/333,573
FILING DATE: OI November 1994
ATTORNEY/AGGNT INFORMATION:
NAME: Maria C.H. Lin
RECISTRATION NUMBER: 29,323
REFERENCE/OCCKET NUMBER: 1151-4118PC
TELECOMMUNICATION INFORMATION:
TELECOMUNICATION INFORMATION:
TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Gaps
                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                Query Match 92.0%; Score 138; DB 5; Length 61; Best Local Similarity 89.3%; Pred. No. 6e-13; Matches 25; Conservative 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%; Score 138; DB 5; Length 61; 89.3%; Pred. No. 6e-13; 1. Indels 1.ve 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
PCT-US95-13660-3
SCHOOLE 3, Application PC/TUS9513660
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (212)758 4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acids
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 89.3 Matches 25; Conservative
                                                                                               / TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

Search completed: August 7, 2003, 11:23:53 Job time : 11.6364 secs

Fri 'Aug

us-09-491-146a-31.rapb

```
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DOLIVET, MICHEL
APPLICANT: DALON, PRANCOIS
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LADAVIERE, LAURENT
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 1999-09-07.
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
LENGTH: 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: Of Hepatilis Virus
US-09-929-955-2

US-09-811-138-76

US-09-851-138-16

US-09-851-138-16

US-09-851-138-18

US-09-851-138-18

US-09-899-046-50

US-09-899-046-52

US-09-899-046-54

US-09-878-281-54

US-09-878-281-54

US-09-878-281-54

US-09-878-281-54

US-09-973-025-50

US-09-995-808-50

US-09-995-808-50

US-09-995-808-50

US-09-995-808-50

US-09-995-808-50

US-09-955-575-40

US-09-952-575-40

US-09-952-575-40

US-09-952-575-40

US-09-952-575-40

US-09-952-575-40

US-09-952-572-9

US-09-952-572-9

US-09-952-572-9

US-09-952-955-1

US-09-952-955-1

US-09-952-955-1

US-09-951-804-3

US-09-951-804-3

US-09-951-804-3

US-09-951-804-3

US-09-951-804-3

US-09-951-804-3

US-09-951-804-3

US-09-951-804-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
PAGES: 3392-3396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1, Application US/10367677 Publication No. US20030118604A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; RELEVANT RESIDUES: 2 TO 45 US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                               2894
2985
3011
3011
3011
3011
3012
      US-10-367-677-1
      DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
        Sequence 1, Appli
Sequence 8, Appli
Sequence 17, Appl
Sequence 14, Appl
Sequence 78, Appl
Sequence 6, Appl
Sequence 60, Appl
Sequence 152, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1, Appli
Sequence 10, Appl
                                                                                                                                     August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Published_Applications_Ash.

/ cgn2_6/ptodata1/pubpaa/US07_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US07_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US06_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US06_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US07_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US09_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10B_PUBCOMB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
/ cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep:*
                      5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-756-875-8
US-09-921-397-77
US-09-921-397-77
US-09-851-138-46
US-09-851-138-46
US-09-891-046-152
US-09-899-046-152
US-09-899-046-44
US-09-878-281-42
US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-10-367-677-1
US-09-851-138-10
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                          150
1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                           451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                      GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                      protein search, using sw model
                                                                                                                                                                                                                                                                                                                      0.5
                                                                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                       US-09-491-146A-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0000000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           444
744
91
100
1113
137
1166
1166
1166
1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                              score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              138
1138
1138
1138
1138
1138
1138
1138
                                                                                                      protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                                                                                              Sequence:
                                                                                                                                                                                                                       Title:
Perfect ;
                                                                                                                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ş
```

Sequence 4, Appli Sequence 9, Appli Sequence 1, Appli Sequence 20, Appl

Sequence 3, Appli Sequence 1, Appli Sequence 20, Appl

Sequence Sequence

ALIGNMENTS

Length 44;

DB 15;

Score 138;

Appl Appl Appl Appl Appl Appl

Sequence 52, Sequence 54, Sequence 144, Sequence 50, Sequence 50, Sequence 50, Sequence 23, Sequence 23, Sequence 23, Sequence 40,

Appl

Sequence 12, Sequence 18, Sequence

Sequence

Sequence Sequence Sequence Sequence

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 97;
                                                                                                                                                                                                                 Length 91;
                                                                                                                                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 10;
                                                                                                                                                                                                               Score 138; DB 9;
Pred. No. 2.8e-12;
2; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 10
Pred. No. 3e-12;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E: Suite 701-E Columbia Square 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                          1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-406-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER: 1808-157A TELECOMMUNICATION INFORMATION: TELEPHONE: (202)783-6040
                      60/092,339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 60/092
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                             92.0%;
hilarity 89.3%;
Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.0%;
                                                                                                                                      ; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 89.3%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEC ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: D. C. COUNTRY: U. S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
STREET: 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20004
                                                                                                                                                                                                                                                                                                                                                                                                          US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY:
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTICENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
                                                                                                                                                                                                                                                                    STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                      ö
                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft WOLD 6.0 / ASCII text output
CURRENT APPLICATION DATA:
A APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                    Indels
Similarity 89.3%; Pred. No. 1.3e-12; 25; Conservative 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.0%; Score 138; DB 10;
89.3%; Pred. No. 2.2e-12;
tive 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: KANABRER, PATRICIA A.
REGISTRATION NUMBER: 29,775.
REFERENCE/DOCKET NUMBER: INNS:004
                                                        1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPORKTERNTNRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILE REFERENCE: 14114,0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                        Sequence 10, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1, Application US/09758308 Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 74 amino acids
                                                                                                                                                                                                                                                 APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                              AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 92,0°
Best Local Similarity 89.3°
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO
Best Local Similarity
Matches 25; Conserva
                                                                                                                                                                        US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
                                                        ð
                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
```

ô

```
APPLICANT: MAERTENS, GEERT STUYVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                    polypeptides selected from
the hepatitis C virus and
                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 113;
                                                                                     Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Microsoft Word 6.0 / ASCII text output CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 10;
Pred. No. 3.5e-12;
2; Mismatches 1;
                                                                              Score 138; DB 10;
Pred. No. 3.3e-12;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                        US-01-1397-78

Sequence 78, Application US/09921397

Sequence 78, Application US/09921397

Patent No. US20020151484A1

GENERAL INFORMATION:

TATLE OF INVENTION: SID nucleic acids and pol
TITLE OF INVENTION: pathogenic strain of the
FILE REFERENCE: B4809A - JAZ

CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT APPLICATION NUMBER: EP 00402225.7

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 156

SEQ ID NO 78

LENGTH: 113

LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                                                              1 PKPORKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 46, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                Query Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OF HEPATITIS C VIRUS GENOTYPES
AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                       GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REFERENCE: B46094 - JAZ CURRENT APPLICATION NUMBER: 2004/921,397
CURRENT FILING DATE: 2001-08-03
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138; DB 10;
Pred. No. 3.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 08/836,075
FILING DATE: <CIGNTOWNDS
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEYAGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES
AND THEIR USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 14, Application US/0951138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                  Sequence 77, Application US/09921397
Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
ORGANISM: Hepatitis C virus
09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 89.3
Matches 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: TEXAS
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 6
US-09-851-138-14
             US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
WUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                               New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 11; Length 166;
Pred. No. 5.3e-12;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 11; Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
        Mismatches
                                                1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                      5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTERNINRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.0%; Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US/09/899,046 FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US/09/878,28
                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                             ; Sequence 152, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Sequence 152, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
      25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: protein US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER:
                                                                                                                                                                                                                                                                               TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
US-09-878-281-152
                                                                                                                                                                            US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
        Matches
                                                δ
                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA

ZIP: 77210-4433
COMPUTER: USA
COMPUTER: ENORM:
MEDIUM TYPE: FORM:
MEDIUM TYPE: FLORPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORREY/AGENT INFORMATION:
ANAMER AND AMPRICED BATERIALS
ATTORREY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                        Score 138; DB 10; Length 137;
Pred. No. 4.3e-12;
2; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 10;
Pred. No. 4.3e-12;
FILING DATE: 21 OCT 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 1NNS:004
INFORMATION FOR SEQ 1D NO: 46:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: KAMMERER, PATRICIA A.
RESISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-851-138-60; Sequence 60, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                  LENGTH: 137 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 138 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AGENTS
                                                                                                                                                                                                                                                                                                                                           92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 25; Conserval
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
```

```
New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                         Gaps
                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                             Score 138; DB 11; Length 169; Pred. No. 5.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxi NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTION DATE:
CURRENT APPLICATION DATE:
CURRENT APPLICATION DATE:
COMPANDIATION DATE:
COM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

FILING DATE:

PRIOR APPLICATION DATA:

FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. 5.4e-12;
2; Mismatches 1
                                                                                      2; Mismatches
                                                                                                                                               1 PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                  5 PKPORKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTERNTNRRPQDVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 44, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                        Sequence 42, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                             92.0%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 4
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New TITLE OF INVENTION: GENCINDER OF SEQUENCES: 270 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          / MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: protein US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                          Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                      US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 44, Application US/09899046
Sequence 44, Application US/09899046
Sequence 44, Application US/09899046

"UBICATION NO. US20030008274A1
GENERAL INFORMATION:
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: Qenotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER FRADABLE FORM:
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                  of hepatitis C virus diagnosis, prophylaxis and therapy.
                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 169;
                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylax NUMBER OF SEQUENCES: 270
COMPUTER REDEALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
5.3e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 11;
Pred. No. 5.4e-12;
2; Mismatches 1;
                          Pred. No. 5.3e
; Mismatches
                                                                                                                                         PKPQRKTERNTNRRPQDVRFSGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTERNINRRPODVRFSGGGGIVG 28
                                                                                                                                                                                                                                                                                                             ; Sequence 42, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                     .;
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 44: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match 92.0%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                       88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENCTH: 169 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              , MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                          Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                        US-09-899-046-42
                                                                                                                                                                                                                                                       RESULT 12
                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
```

Query Match 92.0%; Score 138; DB 11; Length 169; Best Local Similarity 89.3%; Pred. No. 5.4e-12; Matches 25; Conservative 2; Mismatches 1; Indels C ò

ö

0; Gaps

qq

Search completed: August 7, 2003, 12:01:14 Job time : 14.3636 secs

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Minimum DB Maximum DB

Database

```
Non-A, Non-B Hepat
Non-A, Non-B Hepat
Hepatitis C virus
Sequence of peptid
Fragment of open r
Hepatitis C virus
HCV capsid peptide
HCV capsid virus
Hepatitis C virus
Hepatitis C virus
Encoded by Hepatit
Encoded by Hepatit
NC mosaic protein
Blood transmiscibl
HCV-SI full-length
                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                        Human hepatitis C
Human hepatitis C
HCV core-envelope
HCV core-envelope
HCV core-envelope
HCV core-envelope
HCV core-envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Non-A, Non-B Hepat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide VIIIE base
Anti-HCV antibody
Prototype peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NC mosaic protein amino acid fragment H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                       AAY26952
AAY94410
AAY94409
                                                                                                                                                                                                                                                                                                                                          AAR54065
AAW37380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR69545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY15425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY06680 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Fields HA, Khudyakov YE;
 470
30010
3010
3010
3010
400
400
101
101
101
300
300
300
300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO9910506-A1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-MAR-1999.
AAY06680;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
 NC mosaic protein
NC wosaic protein
HCV type 3 capsid
Hepatitis C virus
Encoded by Hepatiti
                                                                                                           Search time 38.5455 Seconds (without alignments)
115.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             / / SIDSI//ggdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/ggdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/ggdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1985.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1985.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1981.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1991.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1995.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1995.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1995.DAT:*
/ SIDSI/gcgdata/geneseqp-embl/AA1999.DAT:*
/ SIDSI/gcgdata/geneseqpeneseqp-embl/AA2001.DAT:*
/ SIDSI/gcgdata/geneseqpeneseqp-embl/AA2001.DAT:*
/ SIDSI/gcgdata/geneseqpeneseqp-embl/AA2001.DAT:*
/ SIDSI/gcgdata/geneseqpeneseqp-embl/AA2001.DAT:*
/ SIDSI/gcgdata/geneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqpeneseqp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.
                                                                                                                                                                                                                                                                                                                      1107863
                GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                       tal number of hits satisfying chosen parameters:
                                                                                                                                                                                           151
1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                         1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                              August 7, 2003, 11:05:37;
                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY06680
AAY06673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB71258
AAR92968
AAR92971
AAR92971
AAR92970
                                                                                                                                                                                                                                                         Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                         seq length: 0
seq length: 200000000
                                                                                                                                                                            US-09-491-146A-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20
220
23
117
117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
28
189
191
191
191
191
191
470
                                                                                                                                                                                                                                            BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88.7
88.7
88.7
88.7
88.7
88.1
```

Score

8

459786

Result

```
HCV type 3 capsid protein fragment.
                                                                                                                                                                                                                                                    AAB71258 standard; protein; 189 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 2; 30pp; German.
                                                                                                        88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11-FEB-2002; 2002WO-AT00046.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-FEB-2001; 2001AT-0000272.
                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                     Ouery Match
Best Local Similarity 96.23
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis c virus.
                                                                               28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200266621-A1.
                                                                                                                                                                                                                                                                                                           18-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29-AUG-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Heinz FX,
                                                                                                                                                                                                                                                                                AAB71258;
                                                                               Sequence
                                                                                                                                                                                                                              RESULT 3
8$$$$$$$$
                                                                                                                                                             à
                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                  ö
                                                                                      provided. The method is designated restriction endonuclease assisted are useful (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAY06673-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                    The invention relates to a mosalc protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                Score 151; DB 20;
Pred. No. 9.2e-15;
Mismatches 0;
                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NC mosaic protein amino acid fragment A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                               AAY06673 standard; Protein; 28
                                                                                                                                                                                                                                                                100.0%;
           Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97US-0921887.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                uery Match 100.

St Local Similarity 100.

atches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-204671/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                       28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Fields HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY06673;
                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                      RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                              g
```

```
comparising a flavivinus mutant that has a deletion of at least 4 comparising a flavivinus mutant that has a deletion of at least 4 comparising a flavivinus mutant that has a deletion of at least 4 correction acids from the capsid protein, provided that the c-terminal hydrophobic region is not affected by the deletion. The vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a vide range of flavivirus diseases, e.g. yellow fever, Japanese encephalitis, dengue, classical swine fever, bovine viral diarrhoea and hepatitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does not effect immune responses to important proteins; and (iii) can not generate a non-natural virus by recombination. The mutant viruses eliminate the need to produce large amounts of infectious/virulent viruses, and can be produced with less expense. The protective response to flusivalvirus lasts significantly longer than that to killed vaccines. This sequence represents a fragment of the capsid protein from Heptitis c virus (HCV) type I described in the disclosure of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAY0673-681 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids from the capsid protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This invention describes a novel attenuated flavivirus live vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Capsid protein; attenuated vaccine; virucide; antinflammatory; hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 134; DB 20; Length 2 Pred. No. 2.7e-12; 0; Mismatches i; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRPQRKTKRNTIRRPQDVKFPGGGQI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
```

Matches

a

g

AAR92968 XX XX AC AAR9 XX DT 02-C XX

```
envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                   core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunishing against HCV infection. The proteins may also be used to defect antibodies against HCV in serum, sellva, lymphocytes or other monouclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                               envelope 1 and core proteins - as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 134; DB 17; Length 191;
Pred. No. 2e-11;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus isolate DK12 core protein.
                                 Hepatitis C virus isolate S52 core protein.
                                                                                                                                                                                                                                                                              US DEPT HEALTH & HUMAN SERVICES. US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKPORKTKRNTIRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                           DNA and amino acid sequence of HCV used to determine HCV genotype and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR92971 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 4; Page 208; 340pp; English
                                                                                                                                                                                                                                                                                                                                 Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 88.7%;
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95WO-US10398
                                                                                                                                                                                                                95WO-US10398
                                                                                                                                                                                                                                              94US-0290665
02-OCT-1996 (first entry)
                                                                 ij
                                                                                                                                                                                                                                                                                                                               Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                               WPI; 1996-139709/14
                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         191 AA;
                                                                   envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAT16643
                                                                                                                                                                                                                                            15-AUG-1994;
                                                                                                                                                                                                                15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9605315-A2
                                                                                                                                               WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02-0CT-1996
                                                                                                                                                                               22-FEB-1996
                                                                                    hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR92971;
                                                                 E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR92971
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                              envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. Of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection
                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 134; DB 17; Length 191; Pred. No. 2e-11;
                                   Length 189;
                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                 DB 23;
                                 Score 134; DB 2;
Pred. No. 2e-11;
                                                                                                                                                                                                                                                                                                             Hepatitis C virus isolate HK10 core protein.
                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 PKPQRKTKRNTIRRPQDVKFPGGGQI 29
                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR92969 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 4; Page 207; 340pp; English.
                                                                                                                                                                                                            AAR92968 standard; Protein; 191
                                88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Purcell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Match 88.7%;
Local Similarity 96.2%;
les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95WO-US10398
                                           Local Similarity 96.2 nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1996-139709/14
                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA;
189 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAT16642
                                                                                                                                                                                                                                                                                                                                                                                                                          WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-AUG-1995;
                                                                                                                                                                                                                                                                            02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                              hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
 Sequence
                                                                                                                                                                                                                                            AAR92968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92969
                              Query Match
                                                                                                                                                                                                                                                                                                                                             HCV; E1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
```

Matches

ò 윱

388888**%**8

RESULT 5
AAR92969
ID AAR9
XX
AC AAR9

ö

Gaps

4

```
Misc-difference 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR92970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR92970
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "amino acid in this position is designated X in the specification, but codon usage shows that the only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note- "amino acid in this position is designated X in the specification, but codon usage shows that the only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; subtype; polymerase chain reaction; amplification; PCR; primer; probe; antibody; infection.
                                                                                                                            AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                             DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection \,
                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317
                                                                                                                                                                                                                                                DB 17; Length 191;
                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                               Score 134; DB 17
Pred. No. 2e-11;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 144..149
/label- Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label = Met, Thr, Lys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Val, Ala, Glu, Gly
171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /label- Cys, Arg, Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Val, Ala, Asp, Gly
(USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                           Claim 4; Page 209-210; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label= Met, Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /label= Ser, Arg
Misc-difference 167
                              Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                                                                                           AAR96547 standard; peptide; 319
                                                                                                                                                                                                                                             88.7%;
1larity 96.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label- Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /label- Gly
                                                 WPI; 1996-139709/14. N-PSDB; AAT16645.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 172
                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Key
Misc-difference 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 157
                                                                                                                                                                                                                          191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                  10-MAR-1997
                                                                                                                                                                                                                            Sequence
                            Bukh J,
                                                                                                                                                                                                                                                                                                                                                 ζ
ò
                                                                                                                                                                                                                                                                                                        9
```

```
The sequences AAR96526-R9657B represent novel sequences isolated from hepatitis C virus subtypes different from subtypes 1a-c, 2a-d, 3a-f, 4a-j, 5a and 6a. They esp. from the novel subtypes 1a-c, 2a-d, 3a-f, 3d-d, 5a and 6a. They esp. from the novel subtypes 1d-f, 2a-l, 2k, 2l, 3g, 4k-m, 7a-c or types 9, 10 or 1l. The sequences corresp. to the 5' genome. This sequence represents amino acids 1-317 from the HCV types 7c and 8a isolates VM; the Core/El, NS4 or NS5B regions of the Stance of the control of the contr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine; hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sub:type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus poly:nucleic acid unique to unidentified - used to develop probes and primers for new sub:types and to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88.7%; Score 134; DB 17; 96.2%; Pred. No. 3.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
    Leu, Ile, Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus isolate S2 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ź
                                                                                                                                                                            Val
                                                                                                                                                                                                                                                                /label- Asn, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 25; Fig 3; 150pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR92970 standard; Protein; 191
'label- Phe,
                                                                                        /label- Phe,
                                                                                                                                                                        /label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95WO-EP04155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          95EP-0870076
94EP-0870166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Maertens G, Stuyver L;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1996-251460/25
                                          Misc-difference 177
                                                                                                                                                                                                                  Misc-difference 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   319 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAT27958
                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                     WO9613590-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-OCT-1996
```

S

ö

Gaps

ö

( HSSU)

```
CDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-B. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-040439.
                                                 cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-A. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See AAQ40425-Q40439.
                                                                                                                                                                                                Length 470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; non-A, non-B hepatitis virus; NANBHV; liver disease;
polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                         1,
                                                                                                                                                                                            Score 130; DB 14;
Pred. No. 2e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.1%; Score 130; DB 14; 92.3%; Pred, No. 2e-10;
                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 2e-1
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Encoded by Hepatitis C virus clone JK3-B.
                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                   5 PKPQRQTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 4; Page 26-28; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 4; Page 28-30; 44pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                            Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR34475 standard; Protein; 470 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                          AAR34474 standard; Protein; 470
                                                                                                                                                                                            86.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91JP-0153736
                                                                                                                                                                           Query Match
Best Local Similarity 92.3v
Best Local Similarity 92.3v
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1993-130638/16
                                                                                                                                                           470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAQ40432
                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JP05068562-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23-MAR-1993.
                                                                                                                                                                                                                                                                                                                                                                                                                               AAR34474;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                         RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR34475
ID AAR3
                                                                                                                                                                                                                                                                                                                                                                           AAR34474
ò
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                       AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other monounclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                  DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        y match
Local Similarity 92.3%; Pred, No. 2.9e-11;
les 24; Conservative 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA and cDNA of hepatitis C virus - useful as probes diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Encoded by Hepatitis C virus clone JK3-A
                                                                                                                                                                        US DEPT HEALTH & HUMAN SERVICES. US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5. PKPQRKTKRNTIRRPQDIKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                        Claim 4; Page 209; 340pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR34473 standard; Protein; 470
                                                                                                                                                                                                                               Purcell RH
                                                                                                         95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91JP-0153736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91JP-0153736
                                                                                                                                        94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                               Bukh J, Miller RH,
                                                                                                                                                                                                                                                                WPI; 1996-139709/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1993-130638/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      191 AA;
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                   N-PSDB; AAT16644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAQ40431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JP05068562-A.
                                   WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-MAY-1991;
                                                                                                       15-AUG-1995;
                                                                                                                                        15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-JUL-1993
                                                                      22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-MAR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR34473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
```

Query Match Best Loc Matches

\*

AAR34473 ID: AAR3 XX AC AAR3 RESULT 9

g

à

ö

Gaps

9

```
The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein and the artificial mosaic protein is also useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the infection, especially hepatitis C. The method of synthesizing the spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAV06673-683 represent amino acid sequence of each monomer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NSS region.
                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                      New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species – useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                               Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                         85.4%; Score 129; DB 20; 92.3%; Pred. No. 1.4e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'label Ser, Arg, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Len
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Key Location/Qualifiers
Misc-difference 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /label- Gln, Pro,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'label- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /label= His, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label- Cys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /label= Asp, Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /label= Lys, Arg
Misc-difference 418
                                                                                                               Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Non-A, non-B hepatitis virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 92.3
Matches 24; Conservative
                           WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 311
                                                                                                                                                                                                                                                                                                                                                             28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17-JAN-1995
                                                                                                                                                                                                                                                                                                                                                             Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR53417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR53417
             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                 cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B (9405 nucleotides long) and 14 shorter clones were isolated by PCR amplification, including clone JK3-C. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 130; DB 14; Length 470; Pred. No. 2e-10;
                                                                                                 HCV; non-A, non-B hepatitis virus; NANBHV; liver disease; polymerase chain reaction; diagnostic method.
                                                                                                                                                                                                                                                                                                                                                             DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                      Encoded by Hepatitis C virus clone JK3-C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NC mosaic protein amino acid fragment C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (USSH ) US DEPT HEALTH & HUMAN SERVICES,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 4; Page 30-32; 44pp; Japanese.
                                                                                                                                                                                                                                                                                   (SANW ) SANWA KAGAKU KENKYUSHO CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY06675 standard; Protein; 28 AA.
                                                                                                                                                                                                                                91JP-0153736.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                           91JP-0153736
                                        30-JUL-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Khudyakov YE;
                                                                                                                                                                                                                                                                                                                WPI; 1993-130638/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequences of these
See AAQ40425-Q40439
                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         470 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAQ40433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mosaic protein;
                                                                                                                                                                                                                                                        30-MAY-1991;
                                                                                                                                                                                                                              30-MAY-1991;
                                                                                                                                                                       JP05068562-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-MAR-1999.
               AAR34475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Fields HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY06675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY06675
ID AAY
```

ò a 

```
expression cassette including a first polynucleotide region including a 5' non-coding region (NRI) sequence of an RNA virus and at least an N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3' untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the RNA virus and a third polynucleotide region encoding a reporter cof the virus and a third polynucleotide region encoding a reporter molecule, flanked by first and second polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a moner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. William the expression cassette uncletc acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV-SI full-length polyprotein.
non-B hepatitis (NANBH) virus. The cDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were C100 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      and a promoter operably linked RNA transcription of the
                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to nucleic acid construct which comprises an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        detecting the presence of
                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                        ;
;
                                                                                                                                                                                      Score 129; DB 15;
Pred. No. 2e-09;
                                                                                                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Novel nucleic acid construct useful for virus, comprises an expression cassette to expression cassette for minus strand
                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MOLE-) INST MOLECULAR & CELL BIOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 1; Page 70-81; 81pp; English
                                                                                                                                                                                                                                                                                                                                                                               AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hong
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV-S1 full-length polyprotein
                                                                                                                                                                                      85.4%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-JUL-2001; 2001WO-IL00669.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-JUL-2000; 2000US-220248P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sg,
                                                                                                                                                                                    Query Match 85.4
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-280605/32.
N-PSDB; AAD33038.
                                                                                                                                                      3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (EHRL/) EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Tan YH, Lim SP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200208447-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31-JAN-2002
                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                 AAE20477;
                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                 888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This sequence is encoded by the genome of a blood transmissible non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 - nsed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Blood-transmissible non-A non-B hepatitis virus DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; Page 8-20; 22pp; Japanese.
                                                                                                                                                                                                                                                          Asp
                                                                                                                                                                                                                                                                                                                                                              Tyr
                                                                                                                                                                                                                                                                                                                                                                                               Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Met
                                                                                                                    Lys
                                                                                                                                                    Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Thr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lea
              /label- Gly,
                                                                                                                                                    abel Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       abel= Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label- Tyr,
                                                                                                                                                                                    label- Ser,
                                                                                                                                                                                                                        label- Thr,
                                                                                                                                                                                                                                                        label- Glu,
                                                                                                                                                                                                                                                                                                                                                             label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                             label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 abel- Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 abel- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       abel- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .abel- Val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .abel- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .abel- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label- Leu,
                                                label- Asn,
                                                                                                                                                                                                                                                                                         label- Met,
                                                                                                                                                                                                                                                                                                                           abel- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                               abel- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 detection of hepatitis virus
                                                                                                                  abel- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Ser,
                                                                                  label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92JP-0051885
                                                                                                                                                                                                                                                                                                       Misc-difference 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 2488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 2968
                                                                                                                                                                                                    Misc-difference 1036
                                                                                                                                                                                                                                                                                                                                         Misc-difference 1255
                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 2989
                                                                                                                                                                                                                                                                       Misc-difference 1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 2862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1994-163130/20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAQ63499
                                                                                                                                                                  Misc-difference
                              Misc-difference
                                                                Misc-difference
                                                                                                Misc-difference
                                                                                                                                Misc-difference
                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JP06105690-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-APR-1994.
```

us-09-491-146a-30.rag

```
ö
                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is that of an antigen used in a novel assay for determining anti-HCV antibody titre. A sample is diluted to an absorbance of not more than 2.0 and the peptide antigen is added. The method is useful for monitoring the effect of interferon treatment on anti-HCV antibody titres.
                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                    85.4%; Score 129; DB 23; Length 3010; dlarity 92.3%; Pred. No. 2e-09; Conservative 0; Mismatches 2; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Determination of the antibody titre against hepatitis C virus antigen - using dilute solution and HCV antigenic peptide(s)
                                                                                                                                                                                                                                                                                         HCV; non-A, non-B hepatitis virus; antigen; immunoassay;
interferon treatment; monitoring; antibody titre; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match

84.8%; Score 128; DB 16; Length 30;
Best Local Similarity 92.3%; Pred. No. 2.1e-11;
Matches 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                 Hepatitis C virus core antigenic peptide.
                                                                                     AAR84558 standard; peptide; 30 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 1; Page 10; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                94JP-0082160.
                                                                                                                                                                                                                                                                                                                                                                                                                                           94JP-0082160.
                                                                                                                                                                                                                                      06-JUN-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
                                 Ouery Match
Best Local Similarity
Matches 24; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1995-386063/50.
           3010 AA;
                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                           16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                           JP07260792-A
                                                                                                                                                                                                                                                                                                                                                                                     13-OCT-1995.
           Sequence
                                                                                                                                                                                                             AAR84558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                       RESULT 15
AAR84558
ID AAR84
XX
AC AAR84
XX
XX
YX
YY
X S
                                                                                                                g
                                                                                                                                                                                                                                                                                              S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

Search completed: August 7, 2003, 11:14:09 Job time : 39.6364 secs

```
GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 11:05:41; Search time 9.54545 Seconds

(without alignments)
282.095 Million cell updates/sec

Title: US-09-491-146A-30

Perfect score: 151
Sequence: 1 PKPORKTKRNTIRRPODVKPPGGGVIXV 28

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

'tal number of hits satisfying chosen parameters: 283308
```

Database: PIR\_76:\*

1: pirl:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Ainimum DB seq length: 0
Maximum DB seq length: 2000000000

SUMMARIES

	Description	qenome polyprotein							genome polyprotein													genome polyprotein				genome polyprotein				genome polyprotein
SUMMARIES	ID	S41359	S41358	S41360 ·	PC2061	S41288	S41361	S19875	S41356	PC2060	A45573	S21336	S41353	S41355	S41357	S41348	S41371	S41341	S41370	S41369	S41368	S41342	S41344	S41350	35	7	S41347	541343	541346	PQ0393
	DB	7			~	7	~	~	~	~	-	7	7	~	~	~	~	~	~	~	7	~	~	~	N	~	7	~	~	~
	Length	114	114	124	411	492	123	782	108	411	3010	88	108	108	108	108	112	112	114	114		115	115	115	115	115	115	115	118	266
æ	Query	88.7	88.7	88.7	88.7	88.7	86.1	86.1	85.4	85.4	85.4	84.8	84.1	.84	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1
	Score	134	134	134	134	134	130	130	129	129	129	128	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127	127
	Result No.	-	7	m	4	S	9	7	œ	σ	10	11	12	. 13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

88.7%; Score 134; DB 2; Length 114;

Query Match

genome polyprotein genome polyprotein genome polyprotein polyprotein hepa polyprotein hepa polyprotein genome polyprotein genome polyprotein hypotheiral prote genome polyprotein polyprotein polyprotein polyprotein genome polyprotein	RESULT 1 S41359 genome polyprotein - hepatitis C virus (genotype 3, N2) (fragment) Groontains: core protein Grispecies: hepatitis C virus Grispecies: hepatitis C virus Grispecies: hepatitis C virus Grispecies: hepatitis C virus A;Variety: genotype 3, N2 Grispecies: hepatitis C virus A;Variety: genotype 3, N2 Grispecies: hepatitis C virus genotypes 1 to 5 by LiPa. A;Recestion: Analysis of hepatitis C virus genotypes 1 to 5 by LiPa. A;Recestion: S41359 A;Accession: S41359 A;Accession	d protein; core protein; polyprotein core protein *status predicted <mat>  88.7%; Score 134; DB 2; Length 114;  larity 96.2%; Pred. No. 3.9e-12;  Conservative 0; Mismatches 1; Indels 0; Gaps 0;  ORKTKRNTIRRPQDVKFPGGGVI 26                                     </mat>	n - hepatitis C virus (genotype 3, N1) (fragment) protein tis C virus tis C virus pe 3, N1 194 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 194 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 158 158 158 158 159 159 159 159 159 159 159 159 159 159
1748818127E19550	ALIGNMENT us (genot on 26-Jul rouwer, J January 1 C virus :9443886;	prec prec e 13 l. No lisme	(96 numer nu
S21471 S12707 PC1284 J01925 J401926 J401926 J401926 S18031 S18031 J00883 J00883 J00883 J00883 S18030 S18030 S18030 S18030	ALIGNMER  tis C virus (genc is ce_revision 26-Ju G.E.M.; Brouwer, Library, January hepatitis C virus hepatitis C virus heyes 3, N2 virus genome poly	protein; polatus predict Score 134; Pred. No. 3 0; Mismatch FFGGGVI 26	is c virus (genis) ce_revision 26-J G.E.M.; Brouwer, Library, January hepatitis c viru  RA 9461; NID:944388. Trype 3 NI virus genome policore protein; pi core protein; pi
ааааааааааааа	rev Trar ati	# sta	rev. F.M. brar. patil
369 4469 513 520 523 782 782 782 782 782 3010 3010	hepatitis C virus t virus t virus 3, N2 #sequence_revision lleter, G.E.M.; Brou lib Data Library, Jan sis of hepatitis C S41341 comic RNA AN> EMBL.229462; NID:94 e: genotype 3, N2 titis C virus genome	psid protein; core protein t: core protein #status pr 88.7%; Score 1milarity 96.2%; Pred.; Conservative 0; Mis. PKPORKTKRNTIRRPODVKFGGGVI	hepatitis C virus tein C tein C virus 3, NI #sequence_revision leter, G.E.M.; Brou lib Data Library, Jan sis of hepatitis C S41341 omic RNA EMBL; 229461; NID:94 itis C virus genome rotein; core protei
	- he octed 3, 3, 4 set let let SEL SA1)	TER ISEN	he h
00000000000000000000000000000000000000	e pro- trisse pro- 1994 1359 1359 1359 1359 1359 1359 1359 1359	capsid prote uct: core pr h Similarity 25; Conserv I PKPORKTKRN 5 PKPQRKTKRN	ein - ein -
11111111111111111111111111111111111111	SULT 1 1359  nome polyprotein - hepatitis C contains: core protein Species: hepatitis C virus Variety: genotype 3, N2 Date: 19-May-1994 #sequence_rev Accession: 541359 van Doorn, L.J.; Kleter, G.E.M bmitted to the EMBL Data Libra Description: Analysis of hepati Reference number: 541341 Accession: 541359 Wolecule type: genomic RNA Residues: 1-114 <van> Cross-references: EMBL:229462; Cross-references: EMBL:229462; Superfamily: hepatitis C virus</van>	Keywords: caps 1-114/Product: Ouery Match Best Local Sim Matches 25; // 1 PK	RESULT 2 841358 genome polyprotein - hepi W;Contains: core protein C;Species: hepatitis C v A;Variety: genotype 3, N C;Sate: 19-May-1994 #seq C;Accession: 19-May-1994 #seq C;Accession: L.J.; Klete: submitted to the EMBL Da A;Beference number: S41358 A;Accession: S41358 A;Molecule type: genomic A;Residues: 1-114 «VAN» A;Cross-reference: EMBL A;Exesidues: 1-114 «VAN» C;Superfamily: hepatitis C;Keywords: capsid prote:
	RESULT 1 S41359 genome polyprotein - hepatitis N; Contains: core protein C; Species: hepatitis C virus A; Variety: genotype 3, N2 C; Date: 19-May-1994 #sequence_3 C; Accession: S41359 R; van Doorn, L.J.; Kleter, G.E submitted to the EMBL Data Libn A; Description: Analysis of hepa A; Reference number: S41341 A; Accession: S41359 A; Molecule type: genomic RNA A; Residues: 1-114 < VVAN> A; Cross-references: EMBL: 22946; A; Experimental source: genotype C; Superfamily: hepatitis C virus	C;Keywor F;1-114/ Ouery Best L Matche Oy	RESULT 2 S41358 genome polyprotein - hepatitis C virus (genotype N; Contains: core protein C; Species: hepatitis C virus A; Variety: genotype 3, N1 C; Species: hepatitis C virus A; Variety: genotype 3, N1 C; bate: 19-May-1994 #sequence_revision 26-Jul-19 C; Accession: 541358 R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. Submitted to the EMBL Data Library, January 1994 A; Description: Analysis of hepatitis C virus gen A; Reference number: S41341 A; Reference number: S41341 A; Accession: S41358 A; Molecule type: genomic RNA A; Residues: 1-114 < VANN A; Residues: 1-114 < VANN A; Cross-references: EMBL: 229461; NID: 9443884; PII C; Superfamily: hepatitis C virus genome polyprotic; Steywords: capsid protein; core protein; polyprof F; 1-114/Product: core protein #status predicted

~

```
protein; nonstructural protein; ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 prote:
C;Species: hepatitis C virus
A;Variety: isolate JK3
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C;Accession: S19875
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C;Accession: S19875
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C;Accession: S19875
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C;Accession: Squence analysis of putative structural regions of Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Experimental source: isolate JK3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
     C;Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 3, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Molecule type: genomic RNA
A;Residues: 1-123 <VAN>
A;Cross-references: EMBL:229464; NID:9443890; PIDN:CAA82602.1; PID:9443891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Cross-references: EMBL:X61592; NID:959482; PIDN:CAA43789.1; PID:959483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                         A Residues: 1-492 <SEE>
A; Cross-references: EMBL:X76918
A; Cross-references: EMBL:X76918
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; envelope protein; nonstr
F;1-191/Product: core protein #status predicted <COR>
F;192-372/Product: envelope protein #status predicted <ENV>
F;373-492/Product: NS1 protein (fragment) #status predicted <NS1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ру
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 123;
                                                                                                                                                                                                                                                                                                                                                                                                                          Length 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    G.E.M.; Brouwer, J.T.
Library, January 1994
hepatitis C virus genotypes 1 to 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A.Experimental source: genotype 3, N4 (5.5) Apperfamily: hepatitis C virus genome polyprotein C.Keywords: capsid protein; core protein; polyprotein F.1-123/Product: core protein #status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                    Score 134; DB 2;
Pred. No. 1.7e-11;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 130; DB 2;
Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein - hepatitis C virus (genotype 3,
                                                           R;Seelig, R. submitted to the EMBL Data Library, December 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 88.5%;
Matches 23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                            A; Reference number: S41288
A; Accession: S41288
A; Molecule type: genomic RNA
A; Residues: 1-492 SEEP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S18029
A; Accession: S19875
A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C, Accession: $41361
R; van Doorn, L.J.; Kleter,
submitted to the EMBL Data
A; Description: Analysis of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 1-782 <HON>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S41361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                င်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: mRNA
A; Residues: 1-411 <LLD3.
A; Cross-references: GB:L12355; NID:g410169; PIDN:AAA20155.1; PID:g410170
C; Cross-references: GB:L12355; NID:g410169; PIDN:AAA20155.1; PID:g410170
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein; F:192-383, Product: envelope protein E; Histus predicted <SPE>
F:384-411, Product: nonstructural protein E2, NRS1 #status predicted <NPE>
F:196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C; Accession: PC2061
R; Li, J.S.: Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
**A:Title: Identification of the third major genotype of hepatitis C virus in France.
**Reference number: PC2060; MUID:94197744; PMID:8147893
                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                           #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Cross-references: EMBL: 229463; NID: 9443888; PIDN: CAA82601.1; PID: 9443889
                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LIPA.
                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                           N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ρλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'n
                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N:Contains: envelope protein El; nonstructural protein E2/NS1
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                            C; Accession: $41360
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NS1 protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88.7%; Score 134; DB 2; Le
96.2%; Pred. No. 1.4e-11;
... wiematches 1;
  Pred. No. 3.9e-12;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A Experimental source: genotype 3, N3 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-124/Product: core protein #status predicted KM1-124/Product
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88.7%; Score 134; DB 2;
llarity 96.2%; Pred. No. 4.3e-12;
Conservative 0; Mismatches 1
                                                                                                                                                                                                                                                                           C virus (genotype 3,
                                                                                                                                                                                                                                                                                                  N:Contains: core protein
C.Species: hepatitis C virus
A:Variety: genotype 3, N3
C:Date: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                   30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          #HITHIHIHIHIHIHIHIHIH | PKPQRKFPGGGQI 30
                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                               PKPQRKTKRNTIRRPQDVKFPGGGQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein N2 - hepatitis C virus
Similarity 96.2%;
25; Conservative
                                                                                                                                                                                                                                                                              polyprotein - hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . .Jecule type: genomic RNA A; Residues: 1-124 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Perence number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
es 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
les 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cession: S41360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ccession: PC2061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
S41288
                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                               S41360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PC2061
                                                                                  õ
                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
```

```
procein Naya; nonstructural protein Nab; nonstructural protein NSS (Species: hepatitis C virus (C) pate: 19-May-2000 #text_change 19-Jan-2001 (C) pate: 23, 39-53, 1992 (C) pate: 19-May-2000 #text_change from a single Japanese carr A; Reference number: A45573; MUDI: 92295714; PMID: 1318627 (C) pate: A57573 (C) pate: A57573 (C) pate: A57573 (C) pate: A57573 (C) pate: A75773 (C) pate: A75773 (C) pate: A75773 (C) pate: A77741 (C) pate: A7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) cispecies: hepatitis C virus (fragment) Cispecies: hepatitis C virus (cispecies: hepatitis C virus C) Species: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C; Accession: S21336

R: Sato, A.

Submitted to the EMBL Data Library, April 1992

A; Description: A sentitive serodiagnosis of hepatitis C virus infection with two A; Reference number: S21336

A; Accession: S21336

A; Accession: S21336

A; Residues: 1-88 < SAT>
A; Cross-references: EMBL: X65548; NID: 959492; PIDN: CAA46517.1; PID: 959493

C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: polyprotein
                                                                                   enome polyprotein - hepatitis C virus (strain JT)
;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 / Match 85.4%; Score 129; DB 1; Length 30 Local Similarity 92.3%; Pred. No. 5.6e-10; nes 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84.88;
92.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepagenome polyprotein N; Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Accession: PC2060
R;Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Bicohem. Bloppys. Res. Commun. 199, 144-1481, 1994
A.Title: Identification of the third major genotype of hepatitis C virus in France.
A;Reference number: PC2060; MuID:94197744; PMID:8147893
A;Accession: PC2060
A;Access
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                 ö
F;1-191/Product: core protein #status predicted <MAT1>
F;192-383/Product: envelope protein 1 #status predicted <MAT2>
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N5
A; Variety: genotype 2, N5
Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ate: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (genotype 2, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85.4%; Score 129; DB 2; Length 411; 88.5%; Pred. No. 7.6e-11; ive 2; Mismatches 1; Indels
                                                                                                                                                                                                  2; Length 782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 2; Length 108;
                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 85.4%; Score 129; DB 2; Best Local Similarity 92.3%; Pred. No. 2e-11; Matches 24; Conservative 0; Mismatches
                                                                                                                                                                                          Score 130; DB 2;
Pred. No. 1e-10;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKPQRKTKRNTCRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein - hepatitis C virus
                                                                                                                                                                                              86.1%;
92.3%;
                                                                                                                                                                                                                                                                     24; Conservative
                                                                                                                                                                                                  Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
nes 23; Conserv
                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

Gaps

ö

ับ

ò

Gaps

ö

polyprotein - hepatitis C virus (genotype 2, N2) (fragment)

```
RESULT 15
S41348
                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Species: hepatitis C virus
A;Variety: genotype 2, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                %.Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA. A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                                                                          5 by
                                                                                                                                                                                                                                                                                                                                             Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 108;
                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                       Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to A;Reference number: S41341
A;Accession: S41353
A;Accession: S41353
A;Accession: S41353
A;Accession: S41353
A;Accession: S41353
                                                                                                                                                                                                                                                                                                                     94.1%; Score 127; DB 2; Lk
92.3%; Pred. No. 3.8e-11;
...omatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 127; DB 2; L
Pred. No. 3.8e-11;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1-108 < VAN>
A; Cross-references: EMBL: 229458
A; Experimental source: genotype 2, N4
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein #status predicted < MAT>
                                                                                                                                                                                                                                       A Experimental source: genotype 2, N2
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capaid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQKKTKRNINRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Cross-references: EMBL: 229460
A; Experimental source: genotype 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84.1%;
92.3%;
                                                                                                                                                                                                                         A; Cross-references: EMBL: 229456
                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 92.3
Atches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match
Nest Local Similarity 92.5.
Nest Local 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                           C; Accession: 541353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C; Accession: 541355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

```
C;Species: hepatifis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  by LiPA
                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 127; DB 2; Length 108;
Pred. No. 3.8e-11;
0: Mismatches 2; Indels
                                                                                                              Length 108;
                                                                               Score 127; DB 2; Length lu-
Pred. No. 3.8e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C, Accession: $41348
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5
C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capsid protein; core protein; polyprotein F;l-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A:Experimental source: genotype 1, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capaid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted KM1-108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
84.1%; Score 127; DE
Best Local Similarity 92.3%; Pred. No. 3.86
Matches 24; Conservative 0; Mismatches
                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7, 2003, 11:21:48
                                                                                                     Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A, Molecule type: genomic RNA
A, Residues: 1108 <VAN>
A, Cross-references: EMBL: 229451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: August
Job time: 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                                            N; Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S41348
```

us-09-491-146a-30.rsp

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-30 151 1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

tal number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		:::	e bo	is c	e bo				e bo		e bo	na t	plen	vira	a ol	ш) s	chia	pien	a ap	omyc	musculu	musculu	teri	rus	abdi	lae	teri	teri	psis	11a	dium				
	ption									ч	၎	4	c	ᆮ	모	'n			spinacia			homo sapien		saccharomyc	Enm	Snu	mycobacteri		caenorhabdi	xenopus lae				-	clostridium
	Description		56	P27959	001404	001403	P27960	P27961	P26663	P26662	P27958	P26660	P26661	P29846	P26664	080361	P09958	P19711	049937	P18123	046837	09ha64	08k9d9	P05747	P23188	09er35	032982	028193	019408	P21573	006045	P95050	10	215	P24131
			£	7	4	×	2	7	×	Æ		9	8	3			z	z		ω	H	z	Д	F	E	ы		z	-1						<b></b>
	£		POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	RK4_TOBAC	FURI_HUMAN	POLG_BVDVN	RK4_SPIOL	CAT3_MAIZE	ACFD_ECOLI	FN3X_HUMAN	Y401_BUCAP	RL29_YEAST	FURI_MOUSE	FN3K_MOUSE	RL4_MYCLE	FURI_BOVIN	YSMK_CAEEL	YB1_XENLA	RL4_MYCBO	RL4_MYCTU	RK4_ARATH	FRA_DROME	BGAL_CLOAB
		111	-	_	-					_	_	_					_	_ _	_	_	_	_	-	_	_	_	_	-	<u>ر</u>	_	_	_	_	_	-
	Query Match Length I		3010	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3011	282	794	3988	293	496	1520	309	219	28	793	309	230	797	298	303	223	223	282	595	897
æ	Query Match	1 1 1 1 1 1 1	85.4	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1	84.1		78.8	35.1	34.4		32.5					31.1				30.5	ö	30.1	6	φ.	•	29.8	29.8
	Score		129	127	127	127	127	127	127	127	127	127	127	120	119	53		50.5	49	49	4	48.5	48	47	4	46.5	46	4	•	45.5	45	45	45	45	4.5
	Result No.		7	7	m	4	'n	9	۲.	80	O	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

Q91713 xenopus lae P47390 mycoplasma	P1933 mus musculu P1933 rattus norv	P42921 bacillus su Q9czx5 mus musculu	P28276 herpes simp Q60769 mus musculu	P30594 rhizopus ol Q8c2b3 mus musculu	P41128 brassica na
CHRD_XENLA	TAU_MOUSE TAU_RAT SVI MYCTII	RL4_BACSU PII1_MOUSE	IE63_HSV2H TNP3_MOUSE	CHS1_RHIOL HDA7_MOUSE	R131_BRANA
н н					н
941	732 751 1041	332	512	828 938	206
29.8 29.5	29.5	29.1	29.1	29.1 29.1	28.8
44.5	44 4 4 6 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	4 4	4 4	4 4 4 4	43.5
34 35	376	39	4 4 1 2 2	4 4 4	45

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Envelope glycoprotein [Contains: Capaid protein C (Core protein) (P22), (GP86) (GP70) (NS1) (Fragment).
Hepatitis C virus (isolate HC-12) (HCV).
Viruses: SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. MEDLINE-9223023; PubMed-1314459; Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTÉNTIAL.

N-LINKED (GLCNAC.)

N-LINKED (GLCNAC.)
   5 PKPQRKTKRNTYRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPRO02522; HCV_capsid.
InterPro; IPRO02521; HCV_capsid.
InterPro; IPRO02521; HCV_core.
InterPro; IPRO02531; HCV_core.
InterPro; IPRO02531; HCV_core.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01559; HCV_MS1; 1.
ProDom; PD186052; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10074; BAA00968.1; -.
                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11111;
                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                            POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                    RESULT 2
POLG HCV42
ID I 01-AUG
ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR AMINOPERIDAGE.
CAPSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN SI,E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI,E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 129; DB 1; Length 3010;
Pred. No. 1.1e-10;
0; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                            HCV_RdRP.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                       HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                        PPO1538; HCV_NS2; 1.
PP02907; HCV_NS3; 1.
PP01006; HCV_NS4a; 1.
PP01001; HCV_NS4b; 1.
PP00501; HCV_NS5b; 1.
PP00571; helicase_C; 1.
PP00998; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2240
2240
2259
2788
326573 M
                                                                                                                                               HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Match 85.4%;
Local Similarity 92.3%;
tes 24; Conservative
                                                                                                                                                                                                                 Pfam; PF01543; HCV_capsid; Pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_env; 1. Pfam; PF01560; HCV_NSI; 1.
IPR002522; H
IPR002521; H
IPR00251; H
IPR002531; H
IPR002531; H
IPR004109; H
IPR00145; H
                                                                                                                                                                                 InterPro; IPR007095;
                                                                                                                                                                                                  IPR007094;
                                                                                                                                             IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00998;
                                                                                                                                                                                                                                                                                                               PF02907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                    InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                  InterPro;
                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                      ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                            Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
```

(POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL) (POTENTIAL)

(POTENTIAL) (POTENTIAL)

ô

Gaps

ö

Length 520;

Matches

q

80

RESULT 3
POLG\_HCVH4
ID POLG\_HC
DT 01-JUL
DT 01-JUL
DT 16-OCT
DE Genome

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nonstructural protein.

REMOVED FROM CAPSID PROTEIN C BY THE CELLUIAR AWINOPEFTIDASE.

1 115 CAPSID PROTEIN C (POTENTIAL).

MATRIX PROTEIN (FOTENTIAL).

2 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).

4 >520 NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein (Contains: Capsid protein ( Core protein ( P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment). Hepatitis C virus (isolate HCV-KF) (HCV). Wirises; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR: J01975; J01925.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002519; HCV_core.
InterPro: IPR002519; HCV_env.
InterPro: IPR01542; HCV_env.
InterPro: IPR0560; HCV_env.
InterPro: IPR0560; HCV_env.
InterPro: IPR0662; HCV_env.
InterPro: IPR06
                                               Score 127; DB 1; Length 52. Pred. No. 3.4e-11;
    520
56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                         01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                      520 AA.
                                                                                                          0; Mismatches
                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                  84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D10687; BAA01529.1; -.
                                                                Query Match 84.19
Best Local Similarity 92.33
Matches 24; Conservative
                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1115
1191
191
200
200
233
305
418
    520 5
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; G]
Transmembrane;
                                                                                                                                                                                                                                                                                POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                                                         01-JUL-1993
01-JUL-1993
    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                               POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
      Sor
                                                                                                                                             ò
                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMB outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license alreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1))
                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                    Hepatítis C'virus (isolate HCV-476) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                           Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein
                      Score 127; DB 1; Length 5.
Pred. No. 3.3e-11;
  943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                               520 AA
                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                      1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D10688; BAA01530.1; -
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01550; HCV_mS1; 1.
Pfam; PF01560; HCV_MS1; 1.
55704 MW;
                                       Sh 84.1%;
l Similarity 92.3%;
24; Conservative
                                                                                                                                                                                                                                                               STANDARD;
  513 AA;
                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1192
3384
3347
2209
2334
4418
4424
449
                                                                                                                                                                                                                                                             POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
```

JOSS BERFFFFF SOCOOD SOCORRE

(POTENTIAL)

(POTENTIAL) (POTENTIAL) (POTENTIAL)

CHAIN CHAIN CHAIN

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-20001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Hepatitis C virus (isolate AC-J5) (HCV).
                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -i- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-i- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMCTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-92230232; PubMed-1314459; Okamoto H., Kural H., Okada S.I., Yamamoto K., Lizuka H., Cokada S.I., Okada S.I., Tanaka T., Fukuda S., Tsuda F., Mishiro S., Full-length sequence of a hepatitis C virus genome having poor monology to reported isolates: comparative study of four distinct Virology 188:331-341(1992).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL, D10075; BAA00969.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR00251; HCV_core.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01589; HCV_env; 1.
Pfam; PF01580; HCV_MN; 1.
Pfam; PF01580; HCV_MN; 1.
PF00m; P186062; HCV_MN; 1.
PF00m; P186062; HCV_MN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane; Nonstructural protein.
INIT MET 1
                                                                                                            Length 520
                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR AMINOPEPTIDASE.
                                                                                                 Score 127; DB 1; Le Pred, No. 3.4e-11;
                                                                        56476 MW; 1D2BD0A6FF27349B
                                                                                                                                               0; Mismatches
                                                                                                                                                                                                      PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                           84.18;
                                                                                                                                                 Conservative
                                                                        520 AA;
                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SECUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11112;
                                                                                                            Query ...
Best Local Similar
                                                                                                                                                                                                                                                                                                              POLG_HCVJ5
P27960;
                                                       NON_TER
SEQUENCE
                   CARBOHYD
                                    CARBOHYD
                                                                                                            Query Match
                                                                                                                                                                                                                                                                                          POLG_HCVJS
ID POLG_H
AC P27960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                           RESULT 5
                                                                                                                                                                                     ò
                                                                                                                                                                                                                  a
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
6enome polyprotein (Contains : Capaid protein C (Core protein); Matrix protein (Envelope protein M); Major envelope protein E; Nonstructural proteins NSI and NS2] (Fragment).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genotypes.;

Virology 188:331-341(1992)

-i- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-i- SUBUNIT THE VIRION OF THIS VIRUS IS A NUCLEOCARSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAFSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIINE-92230232; PubMed-1314459; Okamoto K., Lizuka H., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a heptitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                       (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                            POTENTIAL).
                                                                                                                                                                                                                                                                                   Score 127; DB 1; Length 737;
Pred. No. 5e-11;
                                                                                                                                                                                                                                                       81207 MW; 3AF699D82AD501B1 CRC64;
              N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   737 AA
                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Interpro; IPR00252; HCV_capsid.
Interpro; IPR00251; HCV_core.
Interpro; IPR002519; HCV_cor.
Interpro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10077; BAA00971.1; -
                                                                                                                                                                                                                                                                                                                 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
NON_TER
SEQUENCE
                                           CARBOHYD
CARBOHYD
                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                     Query Match
                            CARBOHYD
                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
   8
```

```
01-502 (Rel. 2), Last annotation update)
15-58P-2003 (Rel. 42, Last annotation update)
16-58P-2003 (Rel. 42, Last annotation Elevation Elevator 
                                                                                                                                                                                                                                      MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-91140699; Pubmed-1847440; Takamizawa A., Mori C., Fujita J., Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                 ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-96235224; PubMcd-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
*Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84.1%; Score 127; DB 1; Length 737; 92.3%; Pred. No. 5e-11;
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 92.3
les 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1992 (Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                            INIT_MET
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
POLG_HCVBK
ID POLG_H
AC P26663
DT 01-AUG
                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWIss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

X MEDLINE-99227646; PubMed-9568991;
X Yan Y., L1 Y., Munshi S., Sardana W., Cole J.L., Sardana M.,
Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;
T Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
Y virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-647(1998).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELACED FUNCTION.
NS3 AND NS5 AMAY PLAY A ROLE IN THE VIRAL FNA REPLICATION.
C-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
C-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                           the viral
                                                                                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; PubMed-8861916;
LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
MOONDAW E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
 phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ro; IPR007095; RNA_pol_DS_PS.
ro; IPR007094; RNA_pol_PSvir.
PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M58335; AAA72945.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002519; H
InterPro; IPR002511; H
InterPro; IPR002518; H
InterPro; IPR004109; H
InterPro; IPR00145; H
InterPro; IPR00146; H
InterPro; IPR00166; H
InterPro; IPR002166; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A38465; GNWVTC.
1A1Q; 25-MAR-98.
1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1GX6; 10-APR-02
1QUV; 26-JUN-00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-APR-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       nterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1087;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1GX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BOHM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PDB;
```

PF01006;

```
POLG_HCVJA
P26662;
                                      TURN
STRAND
SEQUENCE
   TURN
STRAND
                                                                                                                                                                                                 RESULT 8
POLG_HCVJA
                             HELIX
                                                                                                                                                                                                                                        EEEEES
                                                                                                                                                            a
                                                                                                                                         ò
                                                                                                                        CELLULAR AMINOPEPTIDASE.
CAPELD PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.1.E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01006; HCY_NS5a; 1.
Pfam; PF01066; HCY_NS5a; 1.
Probom; PD186062; HCY_NS1; 1.
SMART: SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Car protein; Envelope protein; Holicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                 POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.
N-LINKED (GL
                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1161
1163
1166
1171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               081
                                                                                                                                                                                                                                                 347
1083
1107
1165
1165
1165
1230
234
234
250
305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1093
10095
11001
11004
11120
11120
11135
11139
11168
11168
11168
11172
11168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       031
039
039
050
050
050
050
072
077
086
090
                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                      ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TURN
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
HELIX
                                                                                                                                                                                                CHAIN
                                                                                                                                      CHAIN
                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HELIX
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3. (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROPEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISCUSSION OF SEQUENCE.
MEDLINE-91192160; PubMed-1849488;
Kato N., Hijkata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Ohkoshi S., Shimotohno K.;
"Molecular structure of the Japanese hepatitis C viral genome.";
FEBS Lett. 280:325-328(1991).
                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDIATRS—9108550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                             Score 127; DB 1; Length 30 Pred. No. 2.3e-10; 0; Mismatches 2; Indels
                                                                                                                                                                327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 92...
Best Local 24; Conservative . 0
                                                                                                                                                                                                                                 84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                1202
1204
1688
1188
1187 118
1189 119
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (RNA)(N).
```

```
POLG_HCVH
P27958;
               ٦
                                    'n
                               . d
                                                                                                                                                                                                                                                                                                                                                                                               CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 127; DB 1; Length 3u. Pred. No. 2.3e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                     PEPTO; IPR002521; HCV_COTE.

PEPTO; IPR002521; HCV_COTE.

PEPTO; IPR002519; HCV_ENY.

PEPTO; IPR002519; HCV_NS1.

PEPTO; IPR002518; HCV_NS3.

PEPTO; IPR000745; HCV_NS4.

PEPTO; IPR000746; HCV_NS4.

PEPTO; IPR001460; HCV_NS58.

PEPTO; IPR001569; HCV_RAP.

PEPTO; IPR001569; HGLGASE_C.

PEPTO; IPR001569; HGLGASE_C.

PEPTO; IPR007094; RNA_POI_PSyIr.

N; PF01543; HCV_COTE; 1.

N; PF01543; HCV_COTE; 1.

N; PF01543; HCV_COTE; 1.

N; PF01543; HCV_COTE; 1.
                                             V_capsid.
                                                                                                                                                                                                                                                           pF01006; HCV_NS4a; 1.
pP01001; HCV_NS4b; 1.
pF01506; HCV_NS5a; 1.
pP00271; helicase_C; 1.
pF00998; viral_RdRP; 1.
                                                                                                                                                                                                                                                                                                                  ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 92.3
Marches 24; Conservative
                                 InterPro; IPR001410; Di
InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002519; H
                                                                                                                                                                                                                                                                                                                              SMART; SM00487; DEXDC;
                                                                                         InterPro; IPR002518;
InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                         InterPro; IPR001490;
InterPro; IPR002868;
InterPro; IPR002166;
                                                                               InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
HSSP; P26663; 1JXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                  nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                             Pfam; |
Pfam; |
                                                                                                                                                                                                                                                 Pfam;
Pfam;
                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                 fam;
```

Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions along as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annoctation update)
6enome polyprotein (Contains: Capsid protein C (Core protein) (P22);
6enome polyprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.29.9); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hoop Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -I-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.
-I-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
NS3-NS4A, NS4A-FORMS A SAID NS5A-NSB.
-I-FUNCTION: NS4A-FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
ACTIVATION OF NS3.
-I-FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-I-FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
ESSENTIAL ROLE IN THE VIROS REPLICATION.
-I-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
-I-CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    oligonucleotide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SUBDIAT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-!- PTM: THE STRUCTURAL PROTEINS C. E1 AND E2 ARE PRODUCED BY PROTECLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-!- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
-!- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates.";
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitan! M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657
MEDLINE-98154321; Pubmed-9493270;
PKPQRKTKRNTIRRPQDVKFPGGGVI
                                        of unwinding.";
Structure 6:89-100(1998)
                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
```

us-09-491-146a-30.rsp

```
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 127; DB 1; Pred. No. 2.3e-10; 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84.1%;
ilarity 92.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                    1411
1417
1420
1436
1439
11453
11463
1480
1480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1627 162
1635 163
1640 165
3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                               24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVJ6
ID POLG_HCVJ6
     CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                        TURN
TURN
HELIX
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                  STRAND
TURN
HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                                                                               STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                    HELIX
STRAND
TURN
STRAND
TURN
STRAND
HELIX
TURN
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
        g
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                         ProDom; PD186062; HCV_NS1; 1.
SMART; SMO4BY; DEXDC; 1.
PO1yprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; 10.FLT.MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
CHARGE RELAY SYSTEM (BY SIMIL CHARGE RELAY SYSTEM (BY SIMIL CHARGE RELAY SYSTEM (BY SIMIL ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC...) (POTEN N-LINKED (GLCNAC...)) (POTEN N-LINKED (GLCNAC...) (POTEN N-LINKED (GLCNAC...)) (POTEN N-LIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                       InterPro; IPR002213; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR001218; HCV_NS2.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001266; HCV_NS4b.
InterPro; IPR001656; HCV_RSAp.
InterPro; IPR001656; HCV_RSAp.
InterPro; IPR001659; HRA_POL_DS_PS.
InterPro; IPR00169; RRA_POL_DS_PS.
InterPro; IPR00149; RRA_POL_DS_PS.
InterPro; IPR00149; RRA_POL_PS_PS.
InterPro; IPR00169; HCV_ENP11.
Pfam; PF01543; HCV_ENP11.
Pfam; PF01549; HCV_ENP11.
Pfam; PF01509; HCV_NS1; 1.
Pfam; PF01006; HCV_NS1; 1.
Pfam; PF01006; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF00998; Viral_RAPP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                       HCV_capsid.
HCV_core.
                                  EMBL; M67463; AAA45534.1;
PIR; A36814; GNWVCH.
                                                         PDB; 1HEI; 25-NOV-98.
PDB; 1A1V; 16-FEB-99.
PDB; 1A1R; 17-JUN-98.
MEROPS; 529.001; -.
TRANSFAC; T04155; -.
                                                                                                                         IPR001410; C
IPR00252; H
IPR00251; H
IPR00251; H
IPR00251; H
IPR00169; H
IPR00149; H
IPR00149; H
IPR00146; H
IPR00166; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

ö

Gaps

```
Transferase; RNA-directed RNA polymerase;
W Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Goat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
T INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
T CHAIN 1 115 CAPSID PROTEIN C POTENTIAL).
T CHAIN 192 383 MAJOR ENVELOPE PROTEIN B; POTENTIAL).
T CHAIN 734 1010 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
T CHAIN 1620 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
T CHAIN 1620 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1620 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1631 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1631 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1641 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-50G-1992 (Rel. 23, Last sequence update)
01-60G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N81); Protein P1 Northuctural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Jaolate HC-J8) (HCV).
                                                                                                                                                                                                                                                                                                                                                                            RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 127; DB 1; Length 30 Pred. No. 2.3e-10; 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; F957F5C1A273BE9E CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
Pfam; PF01506; HCV_NSSa; 1.
Pfam; PF00021; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1091 1091
2036 2038
2811 2811
3033 AA; 329165 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84.18;
92.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                         1111
1169
1241
1323
196
209
234
305
                                                                                    SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 පු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                              MEDILINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayumi M.;
Mucheotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
and divergent regions. "I some thin the reported isolates for conserved
T. J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
I. FUNCTION: THE SMALL PROFEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
OF CHARITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presented of the procursor polyprotein, commonly with Asp or Glu in the P6
postition, Cys or Thr in Pl and Ser or Ala in Pl'.
C. II. CATALTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (G922) (GP35); Envelope glycoprotein El (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Proteasses/helicase NS3 (P70) (Hepperivitin)
(EC 3.4.21.98); Nonstructural protein NS5A (P65); Nonstructural protein NS5B (P66) (P70) (RNA directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-16) (HCV).
Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).

SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPREIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_BSvir.
Pfan; PF01543; HCV_capsid; 1.
Pfan; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR002521; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002511; HCV_NS1.
Interpro; IPR004109; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490; HCV_NS4b.
IPR002868; HCV_NS5a.
IPR002166; HCV_RdRP.
IPR001650; Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; P27958; IHEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                        NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF02907;
PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       <u>-</u>
```

ö

Gaps

SIMILARITY). SIMILARITY). SIMILARITY).

(POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL)

(POTENTIAL)

```
MEDLINE-92330206; Pubmed-1314449;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";
Virology 188:102-113(1992).
I- FUNCTION: THE SMALL PROTEINS NS2B, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
SUBUNTT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
         RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or A1a in P1. CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 127; DB 1; Length 3033;
Pred. No. 2.3e-10;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                              .1 N-LINKED (GLCNAC. . .) (POT 330177 MW; 1A173E7E3381FD1A CRC64;
                                       (BY
(BY
                                      CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                      CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                        N-LINKED GLCNAC
                                                                                          ATP (POTENTIAL).
DECH BOX.
                                                                                                                             N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           th 84.1%;
Similarity 92.3%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                      1091 109
2038 203
2359 235
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVTW
                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                      ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
       CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      between the SAISS Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR ANINOPEPTIDAGE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOED PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                   (RNA)(N):
-1-SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
SEQUENCE FROM N.A. MEDILINE-9223023; PubMed-1314459; Okamoto H., Kural K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T. Eukuda S., Tsuda F., Mishiro S.; "Fuluda S." Tsuda F., Mishiro S.; "Fuluda S." Fuluda F., homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002868; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR004109; HCV_NS3.
InterPro; IPR000745; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002531; HCV_NS1
InterPro; IPR002518; HCV_NS2
InterPro; IPR004109; HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          fam; PF01543; HCV_caps1d; 1
fam; PF01542; HCV_core; 1.
fam; PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SM00487; DEXDC; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF02907; PF01006; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ProDom;
SMART; SI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
```

ö

Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensedisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation E. (GP68) (GP70) (NS1); Proteansex/hellosse NS3 (PF0) (Hepacivitin)
16-SEP-2003 (NS1); Proteansex/hellosse NS3 (PF0) (Hepacivitin)
16-SEP-2003 (Rel. 42, Last)
16-S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                                                                      (POTENTIAL).
                                                               (POTENTIAL)
                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3010;
430 N-LINKED (GLCNAC. . .) (POTENT 448 N-LINKED (GLCNAC. . .) (POTENT 540 N-LINKED (GLCNAC. . .) (POTENT 540 N-LINKED (GLCNAC. . .) (POTENT 556 N-LINKED (GLCNAC. . .) (POTENT 623 N-LINKED (GLCNAC. . .) (POTENT 623 N-LINKED (GLCNAC. . .) (POTENT 645 N-LINKED (GLCNAC. . .) (POTENT 7071 N-LINKED (GLCNAC. . .) (POTENT 2077 N-LINKED (GLCNAC. . .) (POTENT 2077 N-LINKED (GLCNAC. . .) (POTENT 2077 N-LINKED (GLCNAC. . .) (POTENT 2078 N-LINKED (GLCNAC. . .) (POTENT 20
                                                                                                                                                                                                                                                                                                                                                                                                              Score 120; DB 1; Lengtn sv. Pred. No. 2.5e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 KPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 KPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                            79.5%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                  3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11104;
                                                                                           540
556
576
623
645
2041
22240
2529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCV1
P26664;
                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                        CARBOHYD
CARBOHYD
         CARBOHYD
                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCV
         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMIL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lib.sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR ANINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E, (POTENTIAL).

MAJOR ENVELOPE PROTEIN NSI, (POTENTIAL).

NONSTRUCTURAL PROTEIN NSI, (POTENTIAL).

PROTEASE, HELICASE NSI (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

NONSTRUCTURAL PROTEIN NSA (POTENTIAL).

RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .) (POTENTIAL)
                                                   SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (BY SI
CHARGE RELAY SYSTEM (BY SI
CHARGE RELAY SYSTEM (BY SI
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC. N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LEADO2166; HCV_NS5a.
LETO; IPR007095; RNA_DOL_DS_PS.
THEPPO; IPR007094; RNA_DOL_DS_PS.
THR, PPO1543; HCV_Caps1d; 1.
The PPO1542; HCV_Caps1d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                 EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral RARP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_env;
HCV_NS1;
HCV_NS2;
HCV_NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002518;
IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                           1N64; 25-FEB-03.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF02907; PF01006; PF01001; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2014
3474
11083
111083
11165
1230
2033
2033
2034
2050
4177
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     etam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
```

DAR WELL BERN WAR AND DAR WAR

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
STRAIN-cv. Bright Yellow 4; TISSUE-Leaf;
STRAIN-cv. Bright Xellow 4; TISSUE-Leaf;
STRAIN-cv. Bright Xellow 4; TISSUE-Leaf;
TObacco chloroplast ribosomal protein L4.";
Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
-- FUNCTION: THIS PROTEIN BINDS DIRECTLY AND SPECIFICALLY TO 23S RRNA
(EX.SIMILARITY). MAY PLAY A ROLE IN PLASTID TRANSCRIPTIONAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nicotiana tabacum (Common tobacco).
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Epermatophyta: Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGULATION.
-1- SUBCELLULAR LOCATION: Chloroplast.
-1- SIMILARITY: BELONGS TO THE L4P FAMILY OF RIBOSOWAL PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FURI_HUMAN STANDARD; PRT; 794 AA. P09958; 014336; 01-MAR-1989 (Rel. 10, Created) 01-MAR-1990 (Rel. 14, Last sequence update) 15-SEP-2003 (Rel. 42, Last annotation update) Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue cleaving enzyme) (PACE) (Dibasic processing enzyme).
                                                                                                                                                                                                                                                                                                                                                                                                                                    30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
50S ribosomal protein 14, chloroplast precursor (R-protein L4).
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; Irruckalo., Pfan. 1. Prosonal_L4; 1. Ribosomal protein; Transit peptide. Ribosomal protein; Trans-binding; Chloroplast; Transit peptide. TRANSIT 1 43 CHLOROPLAST (BY SIMILARITY). CHAIN 4 282 50S RIBOSOMAL PROTEIN L4. DOWAIN 254 282 ASP/GLU-RICH (ACIDIC).
                                                Score 119; DB 1; Length 3011;
Pred. No. 3.5e-09;
1; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 53; DB 1; Length 282;
Pred. No. 1.7;
5; Mismatches 6; Indels
  327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                          282 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKPQRKT --- KRNTIRRPQDVKFPGGGVIY 27
                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                              EMBL, AB010878; BAA31510.1; -. PIR; T01739; T01739. InterPro; IPR002136; Ribosomal_L4/L1E.
                                                                                                                                                                                                                                                                                                                                                                                                        30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last seq
30-MAY-2000 (Rel. 39, Last ann
                                                   78.8%;
84.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ch 35.1%;
1 Similarity 43.3%;
13; Conservative
                                                                               Local Similarity 84.6
es 22; Conservative
                                                                                                                                                                                                                                                                                                                                                          STANDARD;
  3011 AA;
                                                                                                                                                                                                                                                                                                                                                       RK4_TOBAC
080361;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           108
SEQUENCE
                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
FURI_HUMAN
                                                                                                                                                                                                                                                                                                    RESULT 14
                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8844448
8844448
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE NOTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
CHARGE RELAY SYSTEM (I
CHARGE RELAY SYSTEM (I
CHARGE RELAY SYSTEM (I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                              LUCETPO: JERO02166; HCV_NS5a,
LUCETPO: JERO0166; HCV_NGAP.
LINEAPTO: JERO01056; Hellcase_C.
InterPro: JERO07095; RNA_POI_DS_PS.
InterPro: JERO07095; RNA_POI_DS_PS.
JEAM: PF01543; HCV_corp. 1.
Pfam: PF01543; HCV_corp. 1.
Pfam: PF01543; HCV_corp. 1.
Fam: PF01550; HCV_NS1; 1.
Am: PF01550; HCV_NS1; 1.
Am: PF01538; HCV_NS7:
Am: PF01538; HCV_NS7:
Am: PF01539; HCV_NS7:
Am: PF01540; HCV_NS7:
Am: PF015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                           InterPro; IPR00518; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR000745; HCV_NS3.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001868.
                                                                                                                                                      InterPro; IPR002522; HCV_capsid
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            rtam; PF01538; HCV_NS2; 1. Pfam; PF02907; HCV_NS3; 1. Pfam; PF01001; HCV_NS4s; 1. Pfam; PF01001; HCV_NS4s; 1. Pfam; PF01001; HCV_NS5a; 1. Pfam; PF00299; VIRAL_RGRP; 1. Probom; PD186062; HCV_NS1; 1. SNART; SM00487; DFYPO
                                                                                                                                                                                                               InterPro; IPR002519; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       729
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10066
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    645
2041
2077
2240
2364
2789
                   PDB; 1A1V; 16-FEB-99
PDB; 1HEI; 25-NOV-98
                                                                                                                                  InterPro; IPR001410;
                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
```

ä

13

OR PACE OR FUR

```
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBGELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE BNDOPLASMIC RETICCLUM (ER).

SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN.

TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.

LOCALIZATION AND RECYCLING FROM THE CELL SURFACE.

PLOCALIZATION AND RECYCLING FROM THE CELL SURFACE.

PLOCALIZATION AND RECYCLING FROM THE CELL SURFACE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 296-794 FROM N.A.
MEDLINE-87053858; PubMed-3023061;
Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C.,
Bloemers H.P.J., van de Ven W.J.M.;
"Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein.";
EMBO J. 5:2197-2202(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          respective precursors.
COFACTOR: CALCIUM-DEPENDENT.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM (ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD LEAD TO THE ACTIVATION OF FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
"Homology modelling of the catalytic domain of human furin. A model
for the eukaryotic subtilisin-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                  Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                      TISSUE-Blood;
MEDLINE-90175002; PubMed-2408021;
van den Ouweland A.M.W., van Duljnhoven H.L.P., Keizer G.D.,
Dorssers L.C.J., van de Von W.J.M.;
Structural homology between the human fur gene product and the
subtilisin-like protease encoded by yeast KEX2.";
Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-91321735; PubMed-1713771;
Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Klefer M.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [4]
3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-94291619; PubMed-8020465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92332543; PubMed-1629222;
                       Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                 NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROPEPTIDE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Brake A.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROCESSING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <u>:</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <u>+</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00136; SUBTILASE_ASP; 1.
PROSITE; PS00137; SUBTILASE_HIS; 1.
PROSITE; PS00138; SUBTILASE_SER; 1.
Hydrolase; Serine protease; Transmembrane; Glycoprotein; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. . .) (POTENTIAL). (GLCNAC. . .) (POTENTIAL). (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC. .) (POTENTIAL N-LINKED (GLCNAC. .) (POTENTIAL CLEAVAGE (SECOND AUTO-).
CELLATACHMENT SITE (POTENTIAL).
CELL STRACHMENT SITE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 52; DB 1; Length 794; Pred. No. 7.3; 4; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANS GOLGI NETWORK SIGNAL.
                                                                                                                                                                                   GO: GO: 0005794; C:GO1g1 apparatus; TAS.
GO: GO: 0004276; F:furin activity; TAS.
GO: GO: 0004276; F:furin activity; TAS.
GO: GO: 0007267; P:coll-cell signaling; TAS.
GO: GO: 0005608; P:coll-cell signaling; TAS.
InterPro; IPR002804; P-domain.
InterPro; IPR002804; P-domain.
InterPro; IPR00209; Peptidase_S8.
Ffam; PF00082; Peptidase_S8; I.
PFam; PF00082; Peptidase_S8; I.
PRINTS, PR00777; P-domain; I.
Probom; P000777; P-domain; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10C44DD5892EF85D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7, 2003, 11:20:03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                             FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 ORKTKRNTIRRPODVKFP 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MW;
                                                                                 EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -. PIR; A39552; KXHUF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34.48;
50.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    August
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      794 AA;
                                                                                                                                                    MEROPS; S08.071; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: Augus
Job time : 5,90909 secs
                                                                                                                                                                                                                                                                                                                                                                                                     Zymogen; Calcium
SIGNAL 1
                                                                                                                                                                            MIM; 136950;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .,
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
```

hepatitis hepatitis

```
0.084983
0.084976
0.081341
0.081341
0.08893
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.080500
0.08
                   080P81
080P83
080P76
080P86
081340
081341
080M74
068893
068893
088CX2
                                                                                                                                                                                                                                                                                                                                                                                     080MIS
080MI9
080MJ1
080MJ0
080RI7
                                                                                                                                                                                                                      QBBCW9
QBBCW8
                                                                                                                                                                                                                                                      QBBCW7
QBBCW6
QBBCW4
QBQRJ1
QBQMI6
                                                                                                                                                                                                                                                                                                                                        Q8QM18
Q8QRJ4
Q8QRJ5
                                                                                                                                                                                                        O8BCX0
(without alignments)
281.845 Million cell updates/sec
                                                                                                            August 7, 2003, 11:05:41; Search time 25.6364 Seconds
            5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                           tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                               830525 segs, 258052604 residues
                                                                                                                                                                                              151
1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV
          GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                              protein search, using sw model
                                                                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                              US-09-491-146A-30
```

score:

Perfect

Sequence:

protein

Run on:

Scoring table:

Searched:

hepatitis hepatitis

## ALIGNMENTS

sp\_archea:\* sp\_bacteria:\*

SPTREMBL\_23:\*

Database :

hepatitis of hepat

STRAIN-K3a; MEDLINE-95053917; PubMed-7964640; Yamada N., Manihara K., Mizokami M., Ohba K., Takada A., Tsutsumi M., Full length sequence of the genome of hepatitis C virus type 3a:
comparative study with different genotypes.";
J. Gen. Virol. 75:3294-3284(1994).
-1- SUBDUNT: THE VIRION E THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
HSSP: D27958; 1HEI. Genome polyprotein.
Heparitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
NCBI\_TaxID=11103; STRAIN-K3a; Date T.; Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases. Last sequence update) Last annotation update) PRT; 3021 AA Created) HCV\_capsid. HCV\_core. HCV\_env. HCV\_NSI. 01, 01, 23, DEAD. Q81495 PRELIMINARY; Q81495; Q1-NOV-1996 (TIEMBLEEL 0 01-NOV-1996 (TIEMBLEEL 0 01-MAR-2003 (TIEMBLEEL 2 MEROPS; S29.001; -.
MEROPS; U39.001; -.
Interpro; IPRO01410; D
Interpro; IPRO02522; H InterPro; IPR002521; InterPro; IPR002519; InterPro; IPR002531; SEQUENCE FROM N.A. SEQUENCE FROM N.A RESULT 1 Q81495 DDR READ BRANCH DE BRANCH hepatitis chepatitis c hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis Q81495 hepatitis Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. Description 068307 068308 068308 081382 081383 081383 081384 08 SUMMARIES 068308 068310 08JYS2 08JYS3 08JYR4 080P85 Q8QP84 Q8QP75 080P80 sp\_fung1:\*
sp\_human:\*
sp\_lnvertebrate:\*
sp\_mammal:\* sp\_unclassified:\* sp\_vertebrate:\* sp\_rvirus:\*
sp\_bacteriap:\* sp\_organelle:\* sp\_phage:\* sp\_archeap: sp\_rodent:\* sp\_plant:\* sp\_virus:\* Query Match Length DB sp\_mhc:\* 

Score

Š

Result

1000.0 888.7 888.7 888.7 888.7 888.7 888.7 888.7 888.7

```
45 45
45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q68310
Q68310;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     068310
                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOTER REPRENTED DESTRUCTION OF THE SOLUTION OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; U23745; AAA65052.1;
                                                                                                                                                                                                                                                                                                                                                                                                        PROSTTE; PS50507; RDRP_VIRAL; 1.
PROSTTE; PS50511; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3021 AA; 328386 MW; A97418FF36C062A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-HCV-BB21;
Songsivila1 S., Kanistanon D., Kunkitti R.;
*Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3021;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.7%; Score 134; DB 12; Length 45; 96.2%; Pred. No. 3.3e-13; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Score 151; DB 12; 100.0%; Pred. No. 7.4e-14; ative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                         interPro; IPR002166; HCV_RGSa.
InterPro; IPR002166; HCV_RGNB.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Pfam; PP01542; HCV_cars; 1.
Pfam; PP01550; HCV_cnv; 1.
Pfam; PP01550; HCV_NS1; 1.
Pfam; PP01539; HCV_NS2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                               Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186662; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 45
45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                             DEXDc; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
hes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                        ProDom; 1
SMART; S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               968307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S E E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                             virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
LIPOPROTEIN THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
--- SUBMUT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBMUT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL, 1023/48, AAA65055.1; --
Interpro; IPR005222; HCV_capsid.
Pfam; PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                        STRAIN-HCV-BB38;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
STRAIN-HCV-C94009;
STRAIN-HCV-Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.7%; Score 134; DB 12; Length 45; 96.2%; Pred. No. 3.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45 AA; 5014 MW; CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 134; DB 12;
Pred. No. 3.3e-13;
0; Mismatches 1;
                     068308;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence upda
01-OCT-2002 (TrEMBLrel. 22, Last annotation upper protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; U23746; AAA65053.1;
InterPro: IPR02522; HCV_caps1d.
Ppfam; PP01543; HCV_caps1d; 1.
Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88.7%;
ilarity 96.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best_Local Similarity
Matches 25; Conser
```

```
Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                            73
8201 MW;
88.78;
96.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                              88.7%;
96.2%;
 Query Match 88.7
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 88.7
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                  1
73 AA; E
                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                                                               STRAIN-RIG291;
                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein.
NON_TER
NON_TER
SEQUENCE 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        080P85
                                                                                                                          Q8JYR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
Q8QP85
                                                                                                   RESULT 7
QBJYR4
                                                                  qq
                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                STRAIN-RIG253;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases.
-!- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GILYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN EVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AF506612; ARAB3388.1; -
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                     Score 134; DB 12; Length 60;
Pred. No. 4.5e-13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                 ACCE7D9C8B90299A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              B92CCE7D9C8B9029 CRC64;
                                                                                                  01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                             60 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORE PROTEIN.
           PKPORKTKRNTIRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                         EMBL; AF506613; AAM33389.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                     Genome polyprotein (Fragment)
                                                                                                                                                                                                                                                                                                                                                        60
6725 MW;
                                                                                                                                                                                                                                                                                                                                                                                      88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
6881 MW;
                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 96.2
nes 25; Conservative
                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                60 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. STRAIN-RIG256;
                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8JYS3;
Q8JYS3;
01-OCT-2002 (
                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                             Q8JYS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                         RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OBJYS3
ID OE
AC OE
 à
                      g
                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŝ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Magnius L.O.; of hepatitis C virus identified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN EVIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY PROTEIN MAND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AF506621; ARM33397.1; -
InterPro, IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            K
                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                         ö
61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6EDC082DAE8CB796 CRC64;
                                                                                                                                                                                                                                                                     01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUN-2002 (TIEMBLEEL 21, Last sequence update) 101-CTT-2002 (TIEMBLEEL 22, Last annotation update) Core protein (Genome polyprotein) (Fragment). Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 134; DB 12;
Pred. No. 5.6e-13;
0; Mismatches 1;
Score 134; DB 12;
Pred. No. 4.6e-13;
                                                                                                                                                                                                                                  ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100 AA.
                                    0; Mismatches
                                                                                                      56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -JUN-2002 (TrEMBLrel. 21, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI
                                                                                                                                                                                                                                PRT;
```

```
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q8QP71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8QP71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     080P74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
Q8QP71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               080P74
         RATE CCC CCC RATE DOWNER WAS DOWNER WAS DOWNER WAS BUT WAS BUT
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     in St. Petersburg.*;
J. Virol. 76:4034-4043(2002).
1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA (BY SIMILARITY).
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88.7%; Score 134; DB 12; Length 100; 96.2%; Pred. No. 7.8e-13; 1ve 0; Mismatches 1; Indels (
                                                                                                                                                                Length 100;
                                                                                             100
11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (alinina O., Norder H., Mukomolov S., Magnius L.O.;
                                                                                                                                                              Query Match

88.7%; Score 134; DB 12;
Best Local Similarity 96.2%; Pred. No. 7.8e-13;
Matches 25; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                       100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100 AA.
                                                                                                                                                                                                                                                                                 1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-233;
MEDLINE-21904745; Pubmed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=21904745; PubMed=11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AY070178; AAL58592.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
EMBL; AX070180; AAL58594.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 88.7

9est Local Similarity 96.2

4atches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                     Polyprotein.
NON_TER 100
SEQUENCE 100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 10
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8QP72
                                                                                                                                                                                                                                                                                                                                                                                                                       080P87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                 080P87
         DR DR DR SO FT
                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
MEDINE-21904745; pubmed-11907242;
Kalinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenotypic recombinant of hepatitis C virus identified
in St. Petersburg.";
J. Virol. 76:4034-4043(2002)
--- SUBRONIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";
J. Virol. 76:4034-4043(2002).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROFIED ENVELOPE. CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPOTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; ANO70193; AALSB607.1;
-1
Interpro; IPR002522; HCV_capsid;
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OBOP74;
01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 22, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                      Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 100;
                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                  100
11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                 88.7%; Score 134; DB 12; 96.2%; Pred. No. 7.8e-13; 11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 134; DB 12;
Pred. No. 7.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPORKTKRNTIRRPODVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AYO70194; AAL58608.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.7%;
96.2%;
                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 90.4.
The 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                       100 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                  100
                                                                                                                                                                                                                          Polyprotein.
NON_TER 1
SEQUENCE 10
```

```
Ralinina O., Norder H., Mukomolov S., Magnius L.O.;
A natural intergenotypic recombinant of hepatitis C virus identified
J. Virol. 76:4034-403(2002).
-1 SIBBUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AV070188; AALSB602.1;
-InterPro; IPR002522; HCV_capsid.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11289 MW; F3BC90F5F595F18B CRC64;
      01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 134; DB 12;
Pred. No. 7.8e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTIRRPQDVKFPGGGQ1 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                  MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88.7%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 96.2*
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER 100 1
SEQUENCE 100 AA;
                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              080P84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
Q8QP84
      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                            STRAIN-631;
MEDLINE-2190745; PubMed-11907242;
MEDLINE-2190745; PubMed-11907242;
Malinina O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";
J. VITO.1. 76,4034-4043(2002).
-1- SUBGNAT: THE VIRION OF FHIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AV070191; AAL58605.1; -.
InterPro; IPR002522; HCV.capsid.
Pfan; PF01543; HCV.capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kalinina O., Norder H., Mukomolov S., Magnius L.O.;

"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";

J. Virol. 76:4034-403(2002).

-1. SGBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; ANO70185; ALLS8599.1;

InterPro; IPR002232; HCV_capsid.

PFam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                         88.7%; Score 134; DB 12; Length 100; 96.2%; Pred. No. 7.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                         100 100
100 AA; 11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11289 MW; F3BC90F5F595F18B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88.7%; Score 134; DB 12;
96.2%; Pred. No. 7.8e-13;
11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOUN-2002 (TIEMBLIA). 21, Created)
01-JUN-2002 (TIEMBLIA). 21, Last sequence updo
01-CT-2002 (TIEMBLIA). 22, Last annotation updo
CT-CT- protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-21904745; PubMed-11907242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 96.2 les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 96.2 les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE 100 AA;
                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
  Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100
                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         080P80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
080P77
ID 080P7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
Q8QP80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ID DTT DTT DTT SERVICE SERVICE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SO DO SE RESERVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

ö

Length 100;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                            Kalinia O., Norder H., Mukomolov S., Magnius L.O.;
"A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg.";
"J. Virol. 76:4034-4043(2002).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENTELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GENCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN A AND GENCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; AALO70181, AALS8955.1;
-1 INTERPRO: IPRO05522; HCV_capsid.
PÉQUI, PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                   01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match

88.7%; Score 134; DB 12;
Best Local Similarity 96.2%; Pred. No. 7.8e-13;
Matches 25; Conservative 0; Mismatches 1;
100 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                              STRAIN-718;
MEDLINE-21904745; Pubmed-11907242;
```

100 AA.

PRELIMINARY;

**080P77** 

οg

Search completed: August 7, 2003, 11:19:02 Job time: 26.6364 secs

```
Query Match 100.0%; Score 151; DB 3; Best Local Similarity 100.0%; Pred. No. 1.5e-15; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LUPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE. "
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRAGMENT TYPE: internal ORIGINAL SOURCE:
61
61
61
61
61
74
74
74
100
1115
1115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DAŢE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; ORGANISM:
US-08-921-887-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-921-887-30
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                         August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 175,
Sequence 175,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 30,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1
Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence (Sequence (Sequen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /cgn2_6/ptcdata/1/laa/5A_COMB.pep:*
/cgn2_6/ptcdata/1/laa/5B_COMB.pep:*
/cgn2_6/ptcdata/1/laa/6A_COMB.pep:*
/cgn2_6/ptcdata/1/laa/6B_COMB.pep:*
/cgn2_6/ptcdata/1/laa/RDMB.pep:*
/cgn2_6/ptcdata/1/laa/PCTUS_COMB.pep:*
                       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-921-887-30
US-08-290-665A-187
US-08-290-665A-187
US-08-290-665A-190
PCT-US95-10398-188
PCT-US95-10398-188
PCT-US95-10398-189
US-08-836-075A-44
US-08-835-886C-230
US-08-921-887-25
US-08-93-886C-192
US-08-93-886C-192
US-08-93-886C-192
US-08-93-886C-192
US-08-93-886C-192
US-08-93-886C-192
US-08-93-886C-192
US-08-93-886C-192
US-08-31-887-25
US-08-31-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                       protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                   US-09-491-146A-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                   OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ş
```

```
Sequence 30, Application US/08921887
Patent No. 603071
GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONGLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
                                                                                                                                                                                         Sequence Sequence
                                                                                                                                                                                                                                                                                 Sequence Sequence 1
                                                                                                                                                                                                                                                                                                                                                                         Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence S
                                                                                            Sequence
                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 28,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA

ZIP: 30303-1769

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
STATE: GA
                                                                                                                          US-08-262-037-26
PCT-US93-08638-9
PCT-US95-13660-3
US-08-835-075A-10
US-08-635-886C-198
US-08-835-886C-198
US-08-835-886C-232
US-08-835-886C-232
US-08-835-886C-832
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-14
US-08-836-075A-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-244-116B-15
US-08-836-075A-46
US-07-946-054-9
US-08-083-947-23
US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-315-850-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -08-501-195-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFCATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WAREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3790
INFORMATION FOR SEO ID NO: 30:
SEQUENCE CHARATERISTICS:
LENGTH: 28 aming acids
```

Gaps

ö

Indels

à

```
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R.H.
NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88.7%; Score 134; 96.2%; Pred. No. 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/290,665A FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 188, Application US/08290665A
Patent No. 5882852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEFX: 421792
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 antho acids
                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W BORK
REGISTRATION NUMBER: 36,459
REFERENCE/MOCKET NUMBER: 20,
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS
               CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FIN
STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ) individual isolate: HK10 US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: MORGAN CORPET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: BUKH, J., MIL
APPLICANT: PURCELL, R. H.
TITLE OF INVENTION: NUCL
TITLE OF INVENTION: CORE
TITLE OF INVENTION: CORE
TITLE OF INVENTION: AND
TITLE OF INVENTION: SEQUENCES OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA
NUMBER OF SEQUENCES:
                                                                                         NEW YORK
: USA
                                                                          NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE
                                                                                                                        10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
US-08-290-665A-188
                                                                                                          COUNTRY:
                                                                    CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ά
                                                                                                                                                                  APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 28;
                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATION SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                  ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 134; DB 3;
Pred. No. 4.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 4.5e
0; Mismatches
             Sequence 187, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: BUKH, J, MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 23, Application US/08921887
Patent No. 6030771
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SS: not relevant
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein HYDOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                            30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                             Š
                                                                                             US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-921-887-23
                   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qa
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                             AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEAPITITS C VIRUS MID THE USE OF REAGENTS DELIVED FROM THESE. SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 5; Length 191;
                                                                                            Length 191;
                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88.7%; Score 134; DB 5; 96.2%; Pred. No. 3.6e-12;
                                                                                            Score 134; DB 2;
Pred. No. 3.6e-12;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                       R.H. NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILLILILILILILILILILI 1
PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                UMBER: PCT/US95/10398
15-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/29/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                               PCT-US95-10398-187
; Sequence 187, Application PC/TUS9510398
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELERAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 187:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
SOFTWARE: WORDPERFECT 5.1
                                                                                          Query Match
Best Local Similarity 96.2%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HK10
         ); ORGANISM: homosapiens
); INDIVIDUAL ISOLATE: DK12
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 96.2'
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: AMI
TITLE OF INVENTION: AMI
TITLE OF INVENTION: COM
TITLE OF INVENTION: SEQ
NUMBER OF SEQUENCES: 26:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: RICHARD W. BORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET: 345 PARK
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10154
                                                                                                                                                                                                                                                                                            RESULT 6
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEAPITITE C VIRUS MAID THE USE OF REAGENYE DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2; Length 191; 3.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 190, Application US/08290665A

Patent No. 5882852

GENERAL INFORMATION

APPLICANT: BUKH, J., MILLER, R.H. AND

TITLE OF INVENTION: NOCLEOTIDE AND DEDUCED

TITLE OF INVENTION: CORE GENES OF INVENTION: OPE

TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE TITLE OF INVENTION: AMINO ACID SEQUENCES OF TITLE OF INVENTION: AMINO ACID SEQUENCES OF TITLE OF INVENTION: AMINO ACID SEQUENCES IN DIAGNOSTIC MENUMBER OF SEQUENCES: 263
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEFAX: (212) 751-6849
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 antho acids
TYPE: amino acid
TYPE: amino acid
STRANDENESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 88.7%; Score 134; Best Local Similarity 96.2%; Pred. No. 3 Matches 25; Conservative 0; Mismatch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
ATTONNEY,AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/POCKET NUMBER: 36,459
REFERENCE/POCKET NUMBER: 3026-4116
TELECHONE: (212) 758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 421792
INFORMATION FOR SEO ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                          ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                            SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NEW YORK
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: unk
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                 TOPOLOGY: unk
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
US-08-290-665A-190
                                                                                                                                                                                                                                                                                                                                                                                                    US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
```

```
Sequence 44, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
APPLICANT: MARRIENS, GEERT
APPLICANT: STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: ABD THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
AND THE USE OF REAGENTS DERIVED FROM THESE SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 134; DB 5;
Pred. No. 3.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                         UMBER: PCT/US95/10398
15-AUG-1995
                                                                                                                                                                                                                                                                                                                           FILING DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E: ARNOLD, WHITE & DURKEE P.O. BOX 4433
                                                                                                                                                                                                                                                         PC-DOS/MS-DOS
                                                                            MORGAN & FINNEGAN
                                                                                                                                                                                                                                  COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 20,
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 42.192
INFORMATION FOR SEQ ID NO: 190:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                    SOFTWARE: WORDFERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88.7%;
darity 96.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: homosapiens
: INDIVIDUAL ISOLATE: DK12
PCT-US95-10398-190
                                                                      ADDRESSEE: MORGAN & FIN
STREET: 345 PARK AVENUE
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: unknown ORIGINAL SOURCE:
                                   NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                    NEW YORK
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: P.O. BCCITY: HOUSTON STATE: TEXAS
                                                                                                                                                                       10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                    TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE TITLE OF INVENTION: ASQUENCES IN DIAGNOSTIC METHODS AND VACCINES CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R.H.
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88.7%; Score 134; DB 5; Length 191; larity 96.2%; Pred. No. 3.6e-12; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Mismatches
                                                                                                     APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NOCLEOTIDE AND DET TITLE OF INVENTION: CORE GENES OF ISOI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             тимыек: PCT/US95/10398
15-AUG-1995
                                                               Sequence 188, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 190, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10:
                                                                                                                                                                                                                                                                                ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 20;
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 15
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGINAL SOURCE:
                                           PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PCT-US95-10398-188
                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

ô

Gaps

```
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88.7%; Score 134; DB 4; Length 319; 96.2%; Pred. No. 6.4e-12; ive 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 189, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BURH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLECTIDE AND DEDUCE
TITLE OF INVENTION: CORE GENES OF ISOLATION:
TITLE OF INVENTION: SEQUENCES:
TITLE OF INVENTION: SEQUENCES IN DIAGNOS:
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; LOCATION: (233)..(233)
; OTHER INFORMATION: Xaa is any amino acid
US-08-635-886C-230
                                                                                                                                                                                               LOCATION: (171)..(172)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                            LOCATION: (174)..(174)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (177)..(177)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: (232)..(232)
OTHER INFORMATION: Xaa is any amino acid
     ION: (161)..(161) INFORMATION: Xaa is any amino acid
                                                                                                    LOCATION: (167)..(167)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15 AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E: FLOPPY DISK
IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 20.
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (212) 758-4800
(212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E: MORGAN & FIN
345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA ZIP: 10154 COMPUTER READABLE FORM: MEDIUM TYPE: FLOPPY
                                                                           NAME/KEY: MISC_FEATURE
                                                                                                                                                                      NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: MISC_FEATURE LOCATION: (177)..(177)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NEW YORK

NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-290-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
APPLICANT: LEROUX ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Geort
TITLE OF INVENTION: URBUS
TITLE OF INVENTION: URBUS
FILE REFERENCE: 2752-18
FILE REFERENCE: 2752-18
FILE REPERENCE: 2752-11-04
FILE REPERENCE: 2752-11-04
FILE REPERENCE: 2752-11-04
FILE REPERENCE: 2752-11-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88.7%; Score 134; DB 3; Length 319; 96.2%; Pred. No. 6.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                     PRILICATION NUMBER: US/08/836,075A
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAME: RAAMERER: PATTORN ANAMER: 29,775
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 29,775
REFERENCE/CHARACTERISTICS:
LENGTH: 319 amilio acids
TYPE: amilio acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 230, Application US/08635886C Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION: (144)..(144)
OTHER INFORMATION: Xaa is any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 88.7
Best Local Similarity 96.2
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYP
US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 230
LENGTH: 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

Gaps

```
GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: BNDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
                                                                        Gaps
                                                                        ö
                            Score 133; DB 5; Length 191;
Pred. No. 5.1e-12;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 129; DB 3;
Pred. No. 2.4e-12;
0; Mismatches 2
                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                           5 PKPQRKTKRNTIRRPQDIKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQKKTKRNTYRRPQDVKFPGGGGI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
CMPUTER: IBM FC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 191, Application US/08635886C Patent No. 6555114
                                                                                                                                                                                                                                       Sequence 25, Application US/08921887 Patent No. 6030771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RECISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
                        Query Match
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 85.4%;
Best Local Similarity 92.3%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRAGMENT TYPE: 9
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
US-08-635-886C-191
                                                                                                                                                                                                 RESULT 13
US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-921-887-25
                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEDATITIES C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                     Length 191;
                                                                                                                                                                                                                                 Score 133; DB 2;
Pred. No. 5.1e-12;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/10398
PILIO DATE: 15-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION:
                                                                                                                                                                                                                                                                                                                                Sequence 189, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFRENCE/DOCKET NUMBER: 2026-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS
                                                                                                                                                                                                                                 88.1%;
92.3%;
TELEX: 421792
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 345 PARK AVENUE
NEW YORK
NEW YORK
                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 92.39
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 191 amino acids
                                                                        TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                   ORGANISM: homosaplens ; INDIVIDUAL ISOLATE: S2 US-280-665A-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: homosaplens
INDIVIDUAL ISOLATE: S2
PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: N
TITLE OF INVENTION: A
TITLE OF INVENTION: C
TITLE OF INVENTION: A
TITLE OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US95-10398-189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: U
ZIP: 10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                            g
```

us-09-491-146a-30.rai

```
Sequence 192, Application US/08635886C

Sequence 192, Application US/08635886C

Patent No. 6555114

GENERAL INFORMATION:

APPLICANT: LEROUX-NOELS, Geert

APPLICANT: DELEYS, Robert

TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C

TITLE OF INVENTION: VIRUS

TITLE OF INVENTION: VIRUS

FILE REFERENCE: 2752-18

CURRENT APPLICATION NUMBER: US/08/635,886C

CURRENT APPLICATION NUMBER: PT/FE94/03555

PRIOR FILING DATE: 1994-10-28

PRIOR FILING DATE: 1994-10-28

SOFTWARE: PAPLICATION NUMBER: EP 93402718.6

NUMBER OF SEQ ID NOS: 286

SOFTWARE: Patentin version 3.1

SEQ ID NO 192

TYPE: PRIOR FILING DATE: PRIOR FILING DA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 275-21
CURRENT PAPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR REPLICATION NUMBER: PT/EP94/03555
PRIOR PAPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PALENT VERSION 3.1
SEQ ID NO 191
LENGTH: 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 85.4%; Score 129; DB 4; Length 450; Best Local Similarity 92.3%; Pred. No. 5e-11; Matches 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 4; Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 129; DB 4;
Pred. No. 5e-11;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 92.3%;
atches 24; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT; ORGANISM: hepatitis C virus US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
```

1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26 

ò qq Search completed: August 7, 2003, 11:23:52 Job time: 10.6364 secs

Fri Aug

```
US-09-878-281-44
US-09-929-955-2
US-09-104-966-2
US-09-10104-966-2
US-09-10104-966-2
US-09-851-138-76
US-09-851-138-76
US-09-899-046-50
US-09-899-046-54
US-09-899-046-54
US-09-899-046-54
US-09-899-046-54
US-09-899-046-54
US-09-899-046-1144
US-09-899-046-1144
US-09-899-046-1144
US-09-899-096-1144
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-995-808-50
US-09-952-575-40
US-09-952-955-1
                                                                                                                                                                                                 US-10-259-275-20
US-09-238-076-2
  August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                         tal number of hits satisfying chosen parameters:
                                                                                  151
1 PKPQRKTKRNTIRRPQDVKFPGGGVIYV 28
                                                                                                                            451899 seqs, 118759770 residues
                                                                                                                                                                                                       Published_Applications_AA:*
                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                   OM protein – protein search, using sw model
                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                            US-09-491-146A-30
```

Title: Perfect score: Sequence:

Run on:

Scoring table:

Searched:

Database :

# ALIGNMENTS

RESULT 1

AGENTS

Appli e 1, Appli > 20, Ar

Sequence Sequence 1

Sequence

Sequence 3 Sequence 1 Sequence 2

Sequence 4, Apr

33, Appl1 176, Appl1 18, Appl1 18, Appl 55, Appl 54, Appl 144, Appl 144, Appl 52, Appl 56, Appl 56, Appl 56, Appl 56, Appl 56, Appl 56, Appl 57, Appl 56, Appl 57, Appl 58, Appl 50, Ap

Sequence Sequence 7

Seguence

Sequence 12, Sequence 18, Sequence 50, Sequence 51, Sequence 52, Sequence 52, Sequence 52, Sequence 50, Sequence 23, Sequence 40, Sequence 50, Seque

```
APPLICANT: MAERIENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                            COMPUTER: IBM PC Compatible
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWHARE: MICTOSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                  CORRESPONDENCE ADDRESS:
RADRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                      Sequence 44, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                               STATE: TEXAS
COUNTRY: USA
ZIP: 7710-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                       NUMBER OF SEQUENCES:
                                                                             US-09-851-138-44
                                                                                                                                                                                                                                                                                                                                                                                                  Appli
                                                                                                                                                                                                                                                                                                                                                Sequence 44, Appl
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                                                                                      Sequence 10, Appli
Sequence 1, Appli
Sequence 71, Appli
Sequence 74, Appl
Sequence 78, Appl
Sequence 78, Appl
Sequence 60, Appl
Sequence 152, App
Sequence 152, Appl
Sequence 42, Appl
Sequence 42, Appl
Sequence 44, Appl
                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
US-09-756-875-8
US-09-921-397-77
US-09-921-397-74
US-09-951-138-46
US-09-851-138-46
US-09-851-138-60
US-09-878-281-152
US-09-878-281-152
US-09-878-281-152
US-09-878-281-152
US-09-878-281-152
US-09-878-281-152
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                              US-09-851-138-44
US-10-367-677-1
US-09-851-138-10
US-09-758-308-1
                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                          Match Length
                                                                                                                                                                                                                                                                                                                                                  319
444
744
1003
1103
1113
1136
1166
1169
                                                                                                                                                                                                                                                                                                             Query
                                                                                                                                                                                                                                                                                                                                                Score
                                                                                                                                                                                                                                                                                                                                                                                                  Result
```

```
Sequence I, Application US/09758308
Patent No. US2002090607a1
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C V
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT FILING DATE: 2001-01-10
PRIOR PILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFWARE: PATENTIN VEFSION 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                         ADDRESSER ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-443
ZIP: 77210-443
ZOMPUTE: Eloppy disk
COMPUTE: Eloppy disk
SOFTWARE: MICTOSOFT WORD 6.0 / ASCII text output
APPLICATION NUMBER: US/09/851,138
FILING DATE: CONNOWND
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 28 JUN 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84.1%; Score 127; DB 10;
1larity 92.3%; Pred. No. 3.8e-11;
Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 127; DB 9;
Pred. No. 4.8e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ch 84.1%;
l Similarity 92.3%;
24; Conservative
                          CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Thes 24; Conservat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                               Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match

84.1%; Score 127; DB 15; Length 44;
Best Local Similarity 92.3%; Pred. No. 2.2e-11;
Matches 24; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                    1; Indels
                                                                                                                                                      Query Match

88.7%; Score 134; DB 10;
Best Local Similarity 96.2%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 1;
                                     TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 44:
                                                                                                                                                                                                                                                            5 PKPQRKTKRNTIRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                     1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 PKPQRKTKRNTNRRPQDVKFPGGGQI 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
SEQUENCE CHARACTERISTICS:
LENGTH: 319 amino acids
                                                                                                                                                                                                                                                                                                                                                                           : Sequence 1, Application US/10367677 'ublication No. US20030118604A1 ENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; RELEVANT RESIDUES: 2 TO 45 US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                         US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 108;
                                                          Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84.1%; Score 127; DB 10;
ilarity 92.3%; Pred. No. 5.8e-11;
Conservative 0; Mismatches 2;
                                                            Score 127; DB 10;
Pred. No. 5.5e-11;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                       18 PKPQRKTKRNTNRRPQDVKFPGGGQI 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                               1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                          Sequence 14, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Sequence 78, Application US/09921397; Patent No. US20020111484A1
GENERAL INFORMATION:
APPLICAMT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                           APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                            AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: P.O. BOX 4433
                                                               84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                            Query Match
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 24; Conserva
                                                                                                                                                                                                                                                   RESULT 7
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
                                                                                                                                                                                        qq
                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                            Sequence 8, Application US/09756875;
Patent No. US20020150990A1;
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TILLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sulte 701-E Columbia Square
STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: PatentIn Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.1%; Score 127; DB 10; 92.3%; Pred. No. 5.2e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: SID nucleic acids and pol
TITLE OF INVENTION: Dathogenic strain of the
TITLE OF INVENTION: pathogenic strain of the
FILE REFERENCE: B4809A - JAZ
CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR PAPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTHARE: PatentIn Ver. 2.1
SEQ ID NO 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808-157A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB 1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/09/756,875 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 77, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 84.1
Best Local Similarity 92.3
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                      STATE: D. C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
                                                                                                                                                                                                                                                                                                              COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-921-397-77
                                                               US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               合
```

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO:
                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: B94870166.9
FILING DATE: CINKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 138;
                       Indels
                   5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match

84.1%; Score 127; DB 10;
Best Local Similarity 92.3%; Pred. No. 7.6e-11;
Matches 24; Conservative 0; Mismatches 2;
                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKPORKTKRNTNRRPODVKFPGGGGI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                    Sequence 60, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 138 amino acids
                                                                                                                                                                                                                                                                                          APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                               LIEVEN
                                                                                                                                                                                                                                                                                                                                                                                     AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
              24; Conservative
                                                                                                                                                                                                                                                                                                                  STUYVER,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-899-046-152
                                                                                                                                                                                                 US-09-851-138-60
              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                           ô
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REPERBNCE: B48094 - JAZ CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT APPLICATION NUMBER: US/09/921,397

PRIOR FILING DATE: 2001-08-02

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 156

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 78

LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUGTON
STATE: TEXAS
COUNTRY: USA
LIP: 77210-4433
COMPUTER READABLE FORM:
WEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
APPLICATION NUMBER: US/09/851,138
FILING DATE: O9-May-2001
PRIOR APPLICATION NUMBER: 09/836,075
FILING DATE: CURKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: CURKNOWN>
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                              Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 137
                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                           DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 127; DB 10;
Pred. No. 7.5e-11;
                                                                                                                                                                                                                                                                                                                                                                      Query Match

84.1%; Score 127; DB 10;
Best Local Similarity 92.3%; Pred. No. 6.1e-11;
Matches 24; Conservative 0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: KAMMEREN, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 46, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.1%;
92.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                              ; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: MAERTENS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

Gaps

ô

Indels

Score 127; DB 11; Pred. No. 9.5e-11; 0; Mismatches

84.1%;

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTKRNTIRRPQDVKFPGGGVI 26
                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                  PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
       Floppy disk
                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                     24; Conservative
                                                                                                                                                                         FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
       MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                          US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy
NUMBER OF SEQUENCES: 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                        Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                          Score 127; DB 11;
Pred. No. 9.3e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 127; DB 11;
Pred. No. 9.3e-11;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                          84.1%; Scc.
92.3%; Pred. No. . .
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKPORKTKRNINRRPODVKFPGGGOI 30
                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQI 30
                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRNTIRRPODVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                  CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
FRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 152, Application US/09878281; Publication No. US20030032005A1
GENERAL INFORMATION: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84.1%;
92.3%;
                                                                                                                                                                                                         LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 92.39
                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.33
Matches 24; Conservative
                                                                                                                                                  FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                            TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  , TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                     US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New sequences of hepatitis C virus
genotypes for diagnosis, prophylaxis and therapy.
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 127; DB 11;
Pred. No. 9.5e-11;
0; Mismatches 2;
1 PKPQRKTKRNTIRRPQDVKFPGGGVI 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKPORKTKRNTNRRPODVKFPGGGOI 30
                                                                                                                              Sequence 44, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-878-281-42

Sequence 42, Application US/09878281

Publication No. US20030032005A1

GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 84.1%;
Best Local Similarity 92.3%;
Matches 24; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
TITLE OF INVENTION: N
TITLE OF INVENTION: 9
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
```

```
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878.281
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U8/362,455
FILING DATE:
COMPUTER: DATA:
APPLICATION OF 42:
CENTRY DESCRIPTION OF 42:
CENTRY DATA:
APPLICATION OF 80 1D NO: 42:
APPLICATION OF 80 1D NO:
```

```
Gencore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 11:05:41; Search time 9.54545 Seconds

(without alignments)
282.095 Million cell updates/sec

Title: US-09-491-146A-29
Perfect score: 154
Sequence: 1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
```

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

al number of hits satisfying chosen parameters:

283308 seqs, 96168682 residues

Searched:

283308

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : PIR\_76:\*
1: pirl:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

							•																								
		Description	genome polyprotein							genome polyprotein										genome polyprotein					•	genome polyprotein					genome polyprotein
SOFETANIES		QI	GNWVC3	S41353	S41355	S41357	S41348	S41371	S41341	S41370	S41369	S41368	S41342	S41344	S41350	S41354	S41345	S41347	S41343	S41346	S21471	S12707	PC1284	JQ1925	JQ1926	JH0711	S19876	\$18031	S18032	PN0677	JQ0883
		80	7	~	N	~	~	7	7	~	7	7	7	7	~	~	~	7	7	~	~	7	C3	~	~	~	7	N	~	7	7
		Match Length DB	3011	108	108	108	108	112	112	114	114	114	115	115	115	115	115	115	115	118	369	441	513	520	523	550	782	782	782	787	874
æ	Query	Match	95.5	9.	σ.	δ.	6	•	ď.	φ.	φ.	φ.	δ.		9.68		6	89.6	6	6	89.6	φ.	φ.		o.	6	ö	σ.		9.68	9.68
		Score	147	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138	138
	Result	S S S	7	7	m	4	S	9	7	ω	σ.	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53

genome polyprotein polypeptide - hepa	genome polyprotein genome polyprotein												
JQ0881 PC2219	GNWVTC	S18030	GNWVCH	S40770	JC5620	JQ1303	GNWVJ8	S41351	S41349	J01584	S41352	521336	S41367
00		-	-	-	-	~	-	~	~	7	7	~	7
874	3010	3010	3011	3011	3014	3033	3033	115	115	640	125	88	109
89.6	89.0 89.0	9.68	9.6	9.68	9.68	89.6	9.6	0.68	89.0	89.0	87.0	86.4	86.4
138	138 138	138	138	138	138	138	138	137	137	1.37	134	133	133

# ALIGNMENTS

RESULT 1
unwo.s general polyprotein - hepatitis C virus (strain HCV-1) N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nons: protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus C;Date: 30-Sep-1992 *Bsequence_revision 30-Sep-1992 *text_change 19-Jan-2001 C;Accession: A39166: P00403: Pon404
R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C. Proc. Natl. Acad. Sci. U.S. A. 88, 2451-2455, 1991
A; Title: Genetic organization and diversity of the hepatitis C virus. A; Reference number: A39166; MUID:91172826; PMID:1848704
A.Accession: A39100 A.Molecule type: mRNA B.Bosidhos: 1-3011 /run
A. Cross-rades: Jost A. C.
J. Gen. Virol. 73, 1131-1141, 1992 A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship
A; Reference number: P00393; MUID:92268871; PMID:1316939
A; Molecule type: genomic RNA
A;Cross-references: DDBJ:D10128
A:Experimental source: isolates E-b16
n,Accession: Puvtu4 A;Status: preliminary
A; Molecule type: genomic RNA
A; Experimental source: isolates E-b17
C; Superfamily: hepatitis C virus genome polyprotein
7:1-115/Product: capsid protein C #status predicted <cpc></cpc>
F:116-191/Product: envelope protein M *status predicted <epm> F:192-389/Product: major envelope protein P *status predicted <wff.< td=""></wff.<></epm>
F:390-729/Product: nonstructural protein NS1 #status predicted <ns1></ns1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <ns2></ns2>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <n4a></n4a>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <n4b> F;2014-3011/Product: nonstructural protein NS5 #status predicted <ns5></ns5></n4b>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,20
Query Match 95.5%; Score 147; DB 1; Length 3011;
ative 2
Qy 1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28

virus genotypes 1 to 5 by LiPA.

ö

```
Analysis of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.68;
                                   A; Reference number: S41341
A; Accession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S41371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S41348
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                        genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
N;Contains: core protein
N;Contains: core protein
S;Species: hepatitis C virus
A;Variety: genotype 2, N2
C;Species: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: $4133
S;Wan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Reference number: $41341
A;Reference number: $41341
A;Reference number: $41353
A;Rolecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N4
C; Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N.Contains: core protein
C;Species: hepatitis C virus
A;Variaty; genorype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41357
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R;van Doorn, L.J.; Kleter, G.B.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                       A. Experimental source: GHBL:229456
A. Experimental source: genotype 2, N2
Superfamily: hepatitis C virus genome polyprotein
sywords: capaid protein; core protein; polyprotein
-108/Product: core protein *status predicted AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Yeerimental source: genotype 2, N4
.pperfamily: hepatitis C virus genome polyprotein
., Acywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 138; DB 2;
Pred. No. 2.4e-12;
3; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 2;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3; Mismatches
    32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQKRNQRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQKKNKRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89.6%;
85.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A, Accession: S41355
A, Molecule type: genomic RNA
Residues: 1-108 <VAN>
A, Cross-references: EMBL: 229458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 85.77
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 4
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
```

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N6
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41348
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lippa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S4131,
S;Wan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
S;Wan Doorn, L.J.; Kleter ilbrary, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5, N5) (fragment)
                                                                                                                                                                                                                                                 Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 108
A.Molecule type: genomic RNA
A.Residues: 1-108 <VAN>
A.Cross-references: EMBL:229460
A.Experimental source: genotype 2, N6
A.Experimental source: genotype 2, N6
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F:1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A:Experimental source: genotype 1, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A: Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                           Score 138; DB 2;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138; DB 2;
Pred. No. 2.4e-12;
3; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 2;
Pred. No. 2.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype
                                                                                                                                                                                                                                                                                                    3; Mismatches
                                                                                                                                                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Match 89.6%;
Local Similarity 85.7%;
les 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: EMBL: 229451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Cross-references: EMBL: 229474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA A; Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA A; Residues: 1-112 <VAN>
```

ö

Length 112;

m

ö

Gaps

ö

Indels

Matches

a

ò

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, NIO
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: $41342
Submitted to the EMBL Data Library, January 1994
A; Reference number: $41341
A; Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A, Variety; genotype 5, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41368
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: 541369
R;van Doorn, L.J; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: 541341
A;Accession: 541363
A;Molecule type: genomic RNA
A;Molecule type: genomic RNA
A;Molecule type: genomic RNA
A;Molecule type: genomic RNA
A;Residues: 1-114 <VVAN>
A;Cross-references: EMBL: 229472; NID: 9443906; PIDN: CAA82610.1; PID: 9443907
A;Experimental source: genotype 5, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A,Molecule type: genomic RNA
A,Residues: 1-114 <VAN>
A,Residues: 1-114 <VAN>
A,Cross references: EMBL:229471; NID:g443904; PIDN:CAA82609.1; PID:g443905
A,Experimental source: genotype 5, N2
C,Superfamily: hepatitis C virus genome polyprotein
C,Superfamily: hepatitis C virus genome polyprotein
C,Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: Z29445; NID:g443852; PIDN:CAA82583.1; PID:g443853
A; Experimental source: genotype 1, N10
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 2;
Pred. No. 2.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89.6%;
llarity 85.7%;
Conservative 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N; Contains: core protein
C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 24; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
           ö
                                                                                                                                                                                                                                                 RESULT 7
341341
Ganda polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S4134 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: Jul: Kleter, G.E.M.; Brouwer, J.T.
R;van Doorn, L.J.; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S41370
genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
N;Contains: core protein
N;Contains: hepatitis C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41370
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the RMBL Data Library, January 1994
.'Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
eference number: S41370
.ccession: S41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Molecule type: genomic RNA
A;Residues: 1-114 <VAN>
A;Residues: 1-114 <VAN>
A;Cross-references: EMBL:229473; NID:g443908; PIDN:CAA82611.1; PID:g443909
A;Experimental source: genotype 5, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  lolecule type: genomic RNA
, Residues: 1-112 <VAN>
A;Cross references: EMBL:229444; NID:g443850; PIDN:CAA82582.1; PID:g443851
A;Experimental source: genotypel, NI
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S41369 genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment) N:Contains: core protein C;Species: hepatitis C virus A;Variety: genotype 5, N3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
           Indels
       1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 2;
Pred. No. 2.5e-12;
           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3; Mismatches
                                                                                                                     1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG
           .,
.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 89.6%;
Best Local Similarity 85.7%;
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.6%;
85.7%;
       24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Simi]
Matches 24; (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ś
```

g ò

ö

Gaps

ö

Indels

RESULT 9

g

à

```
Submone Party Contains: Core protein
C; Species: hepatitis C virus
A; Variety; genctype 2, N3
C; Variety; genctype 2, N3
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541354
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Reference number: S4131
A; Accession: 841354
A; Experimental source: genomic RNA
A; Experimental source: genomic Party C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome
C; Keywords: capsid protein; eore protein; polyprotein
F; 1-115/Product: core protein #status predicted <AMAT>

    hepatitis C virus (genotype 2, N3) (fragment)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search co
Job time
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                       Nicontains: core protein: Servines (genotype 1, N2) (Iragment)
C;Species: hepatitis C virus
A;Variety: genotype 1, N2
C;Species: hepatitis C virus
A;Variety: genotype 1, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41344
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
"mitted to the EMBL Data Library, January 1994
Bsoription: Analysis of hepatitis C virus genotypes 1 to 5 by Lippa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N8
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
A; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
A; A; Reference number: S41341
A; Reference number: S41350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A)Cross-references: EMBL:229447; NID:9443856; PIDN:CAA82585.1; PID:9443857 A)Experimental source: genotype 1, N2 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-115/Product: core protein #status predicted <WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL:229453; NID:g443868; PIDN:CAA82591.1; PID:g443869
A; Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 1, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (genotype 1, N8) (fragment)
                                                           Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89.6%; Score 138; DB 2; Length 115; llarity 85.7%; Pred. No. 2.5e-12; Conservative 3; Mismatches 1; Indels
                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
F;1-115/Product: core protein #status predicted <MAT>
                                                        Score 138; DB 2;
Pred. No. 2.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1
                                                                                                             3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 2.5e
3; Mismatches
                                                                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                               1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.68;
85.78;
                                                        85.7%;
                                               Query Match 89.6
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity
Matches 24; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S41344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Simi
Matches 24;
                                                                                                                                                             ò
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

RESULT 14 S41354

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41345
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Bibmitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Accession: S41345
A;Accession: S41345
A;Molecule type: genomic RNA
A;Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: EMBL:229448; NID:9443858; PIDN:CAA82586.1; PID:9443859
A;Experimental source: genotype 1, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted <MAT>
                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                           Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 115;
1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 138; DB 2;
Pred. No. 2.5e-12;
3; Mismatches 1
                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                           5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7, 2003, 11:21:48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    l Similarity 85.7%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89.68;
                    Query Match 89.6
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     August
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   completed: Augus:
le : 9.54545 secs
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-29

154 1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28 Perfect score: Sequence:

Gapop 10.0 , Gapext 0.5 **BLOSUM62** Scoring table:

127863 seqs, 47026705 residues Searched:

. number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		Description	4 h genome po	hepatitis					h genome po	ᆮ	c	c	ч	h genome po	c	0	c	l nicotiana t	xenopus lae	saccharomyc	2 saccharomyc	lactococcus	) homo sapien					) fusobacteri					3 plasmodium		5 glycine max
		Descr	2666	P27959	001404	001403	2796	P27961	2666	P26662	2795	2666	52666	2984	00026	2865	014738	98036	2157	002354	2550	92×6	79ngn	9epq	399u8	)9ez1	28nws	<b>08rh70</b>	015530	P82593	059721	90691	P2309.	2	P0440
SUMMAKIES		OI	POLG_HCV1							4							_				E						3						ABA		GLC2_SOYBN F
		胃	7	7	-	7	-	7	-	-	-	-	-	~	-	7	-	_	-	-	-	-	~	1	-1	-	-	~	-	~	-	-	-	-	-
		Length	3011	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3010	586	602	282	303	440	979	950	1960	1983	240	240	240	461	556	825	407	339	347	436	485
de	Query	Match						9.6		99.6			89.6	85.1	84.4	35.4	35.4	34.4	34.1			31.8									ö	30.5		30.5	30.5
		Score	147	138	138	138	138	138	138	138	138	138	138	131	130	54.5	54.5		52.5	20	20	49	49	49	4.8	48	48	48	48	48	47.5	47	47	47	4.7
	Result	S S	П	7	m	4	S	9	7	Φ	0	07	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

MEROPS; U39.001; -. InterPro; IPR001410; DEAD. InterPro; IPR002522; HCV\_caps1d.

EMBL; M62321; AAA45676.1; -.

1A1V; 16-FEB-99. 1HEI; 25-NOV-98.

P10723 brugia mala P51816 homo sapien P21711 xenopus lae P08855 oryctolagus	P49428 pichla past O9czx5 mus musculu O9zza0 mus musculu	O551/3 rattus norv Q13625 homo saplen P18123 zea mays (m	QU4U89 saccharomyc P05674 venezuelan
SYN_BRUMA FMR2_HUMAN MIX1_XENLA ICAL_RABIT	UBCX_PICPA PII1_MOUSE PDPK_MOUSE	PDPK_KAT P532_HUMAN CAT3_MAIZE	POLS_EEVV8
ненн.			
548 1311 377 718	332 332 559	1005 496	1254
300.8	2000	2 2 2 2 2 0 0 0 2 0 0 0	29.5
44 46.5 5.5	4 4 4 4 0 0 0 0	45.4 45.5 60.5	45.5
36	2 6 4 4 2 0 0 4	444	4 4 5

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-91172826; PubMed-1848704;

MEDLINE-911728; PubMed-1848704;

MEDLINE-911728; PubMed-1848704;

MEDLINE-911728; PubMed-1848704;

MEDLINE-911728; PubMed-1848704;

MEDLINE-911728; PubMed-1848704;

MEDLINE-911728; PubMed-1848704;

MEDLINE-91172826;

MEDLINE-91172
                                                                                                                           01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annoctation update)
66-One polyprotein (Contains: Capsid protein C (Core protein) (P22);
670-68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
670-68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
670-7: Protease/helicase NS3 (P70) (Hepacivirin)
670-7: Protease/helicase NS3 (P70) (Hepacivirin)
670-7: Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepattis C virus (1solate I) (HCV):
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                    PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or send an email to license@isb-sib.ch).
                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBI_TaxID=11104;
                                                             POLG_HCV1
P26664;
                                    POLG_HCV1
RESULT 1
```

01-AGG-1992 (Rel. 23, Created)
01-AGG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GPG-2003 (GPFO) (NS1) (Fragment).
Hepatitis C virus (isolate HC-J2) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;

513 AA

2

RESULT 2 POLG\_HCVJ2

"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct

MEDLINE-92230232; PubMed-1314459; Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.,

SEQUENCE FROM N.A. NCBI\_TaxID-11111;

```
CELLULAR AMINOPEPTIDASE
CAPSID PROTEIN C (POTENTIAL).
MATIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                      POLYPFOCEALM, DEADC, 1.

COTE PROTECTAL, GLYCOPECTELIN, Transferase, RNA-directed RNA POLYMERASE, Transmembrane, Coat protein, Envelope protein, Helicase, ATP-binding, Transmembrane, Nonstructural protein, Hydrolase, Serine protease; 1815.structure.

INIT_MET. 1. REMOVED FROM CARGAS.
                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Db ;
. 5.4e-13;
0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65F8C9447FCE5AF9
                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 147; DB
Pred. No. 5.4e-
                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                              Proproj IPR007095; RNA_pol_DS_PS.
Proproj IPR007094; RNA_pol_DS_PS.
Proproj IPR007094; RNA_pol_Psvir.
Proproj HCV_core; 1.
PR01542; HCV_core; 1.
PR01539; HCV_env; 1.
PR01539; HCV_NS]; 1.
PR02907; HCV_NS]; 1.
PR02907; HCV_NS]; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95.5%; Scor-
92.9%; Pred
                       HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                           HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
Helicase_C
                                                                                                                                                                                  PF01001; HCV_NS4b; I.
PF01506; HCV_NS5a; I.
PF000271; helicase_C; I.
PF00998; Viral_RGRP; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     327197
                  IPRO02531;
IPRO02518; H
IPRO04109; H
IPRO00745; H
IPRO01490; H
                                                                                                                                                                                                                                  DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2789 278
3011 AA;
                                                                                                                                                                        PF01006;
PF01001;
PF01506;
PF00271;
PF00998;
                                                                                                                                                                                                                                 SM00487;
                                      InterPro;
InterPro;
                                                                 InterPro;
                                                         InterPro;
                                                                           interPro;
                                                                                     InterPro;
                                                                                                         InterPro;
                              InterPro;
                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                     ProDom;
                                                                                                                Pfam; P
Pfam; F
                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                           Pfam;
Pfam;
                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                SMART;
                                                                                                                                                              Pfam;
                                                                                                                                                                        Pfam;
                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genotypes.";
Virology 188:331-341(1992).
-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R EMBL; D10074; BAA00968.1; ...
R InterPro; IPR00252; HCV_capsid.
R InterPro; IPR002521; HCV_capsid.
R InterPro; IPR002521; HCV_core.
R InterPro; IPR002531; HCV_core.
R InterPro; IPR002531; HCV_env.
R InterPro; IPR002531; HCV_env.
R Ffam; PF01542; HCV_env; 1.
R Ffam; PF01560; HCV_env; 1.
R Pfam; PF01560; HCV_NSI; 1.
R Pfam; PF01560; HCV_NSI; 1.
R Pfam; PF01560; HCV_NSI; 1.
R PcDom; PD186062; HCV_NSI; 1.
R PcDom; P0186062; HCV_NSI; 1.
R PcDom; P01860662; HCV_NSI; 1.
R PcDom; P01860664; HCV_NSI; 1.
R PcDom; P01860664; HCV_NSI; 1.
R PcDom; P0186064; HCV_N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
PHERES SOOOS SEED BELL TERESTER SEED BELL SOOO COOCCO COOCCES SEED BELL SEED
```

Conservative

Local Similarity hes 26; Conserv

Best Loca Matches

ö

Gaps

ö

Mismatches

Length 3011; Indels 3

```
520 5
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transmembrane;
                                                                                                                                                                                                                                                                             POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                              POLG_HCVHK
    FT
                                                                                                                                           ò
                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation-the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CELLULAR ANIVORDATIONS.
CAPSID PROTEIN ( FOTENTIAL).
MAJOR ENVELOPE PROTEIN ( FOTENTIAL).
MAJOR ENVELOPE PROTEIN E ( POTENTIAL).
NONSTRUCTURAL PROTEIN BS1/E2 ( POTENTIAL).
N-LINKED (GLCNAC. .) ( POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fullsawa K.;
Genomic characterization and mutation rate of hepatitis C virus
'Genomic characterization and mutation rate of hepatitis C virus
1solated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2799(1992).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GINCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                  01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP6B) (GP70) (NS1); (Frament).
Hepatitis C virus (isolate HCV-476) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein.
REMOVED FROM CAPSID PROTEIN C BY THE
                                                                              ó
                                  Ouery Match 89.6%; Score 138; DB 1; Length 513; Best Local Similarity 85.7%; Pred. No. 1.5e-12; Matches 24; Conservative 3; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coat protein; Envelope protein;
943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                            520 AA
                                                                                                                     1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_capsid.
InterPro; IPR00251; HCV_car.
InterPro; IPR005539; HCV_car.
InterPro; IPR05539; HCV_car.
InterPro; IPR05539; HCV_car.
IPR0Dam; PD186062; HCV_NS1; I.
IPR050m; PD186
55704 MW;
                                                                                                                                                                                                                                                                                               01-JUL-1993 (Rel. 26, Created) 01-JUL-1993 (Rel. 26, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10688; BAA01530.1; -.
                                                                                                                                                                                                                                                          · STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               115
191
191
191
196
196
196
196
196
196
491
4431
4431
513 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=31643;
                                                                                                                                                                                                                                                      POLG_HCVH4
Q01404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                        POLG_HCVH4
                                                                                                                                                                                                                   RESULT 3
SO
                                                                                                                                                        g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nonstructural protein.

REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.

1 115 CAPSID PROTEIN C (POTENTIAL).
6 191 MATRIX PROTEIN (POTENTIAL).
2 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
4 >520 NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
I- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
GENCALOPE GLYCOPTOTEIN EI (GP32) (GP35); Envelope glycoprotein E1 (GP32) (GP68) (GP70) (NS1) (Fragment).
Hepailis C virus (1solate HCV-KF) (HCV).
Viruses; SBRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ( ) (POTENTIAL)
                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .) (POTENTIAL)
                                                                                 Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; J01925; J01925.
InterPro; IPR003522; HCV_capsid.
InterPro; IPR003521; HCV_capsid.
InterPro; IPR003521; HCV_env.
InterPro; IPR003531; HCV_env.
InterPro; IPR003531; HCV_env.
InterPro; IPR003531; HCV_env.
InterPro; IPR003531; HCV_capsid; 1.
Pfam; PF01342; HCV_core; 1.
Pfam; PF01359; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
ProDom; PD186062; HCV_NS1; 1.
ProDom; Glycoprotein; Coat protein; Envelope protein;
520
56499 MW; AA135246CF20D525 CRC64;
                                                       ; DB 1; Len.
. 1.6e-12;
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY.
N-LINKED (GLCNAC.
                                                                              Score 138; DB
Pred. No. 1.6e-
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                             520 AA.
                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                              89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D10687; BAA01529.1; -.
                                                                              Query Match 89.6
Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115
191
191
369
209
234
418
```

```
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11114;
                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJ7
P27961;
                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 6
POLG_HCVJ7
                                                              FT
FT
FT
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial nitities requires a license agreement (See http://www.isb-sib.ch/announce/r send an email to licensedisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           *** SEQUENCE FROM N.A.

*** A MEDLINE-92330232; PubMed-1314459;

*** A Camato H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,

*** A Tanaka T., Fukuda S., Tsuda F., Mishiro S.;

*** Tanaka T., Fukuda S., Tsuda F., Mishiro S.;

*** Tanaka T., Fukuda S., Tsuda F., Mishiro S.;

*** Tonaka T., Fukuda S., Tsuda F., Mishiro S.;

*** Transka T., Fukuda S., Tsuda F., Mishiro S.;

*** Pull-length sequence of a hepatitis C virus genome having poor R. Transka S.;

*** Transka T., Fukuda S., Mishiro S.;

*** Transka T., Fukuda F., Mishiro S.;

** Transka T., Fukuda F., Mishiro S.;

*** Transka T., Fukuda T., Fukuda T., Fukuda A., Fukuda A., Fukuda A., Fukuda A., Fukuda A., Fukuda T., Fuku
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-XUG-1992 (Rel. 23, Created)
01-XUG-1992 (Rel. 23, Last sequence update)
16-CGT-2001 (Rel. 40, Last annotation update)
16-CGT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
Proteins NSI and NS2] (Fragment)
Proteins NSI and NS2] (Fragment)
Hepatitis C virus (isolate HC-J5) (HCV).
Hepacivirus.
                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL) NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                         · · · · (POTENTIAL) · · · · · (POTENTIAL) · · ·
                                      (POTENTIAL)
                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nonstructural protein.
1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                            Length 520;
                                                                                                                                                                                                                                          1; Indels
                                                                                                                       1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                   N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. )
                                                                                                                                                                              Score 138; DB 1;
Pred. No. 1.6e-12;
3; Mismatches 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    737 AA.
                                                                                                                                                                                                                                                                                           1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL, D10075, BAA00969.1, InterPro; IPR002522, HCV. capsid. InterPro; IPR002521, HCV. core. InterPro; IPR002521, HCV. core. InterPro; IPR002521, HCV. core. InterPro; IPR002531, HCV. core. IPEM. PF01543; HCV. core. IPEM. PF01543; HCV. core. IPEM.; PF01539; HCV. core. IPEM.; PF01559; HCV. core. IPEM.; PF01550; HCV. core. IPEM.; PF01560; HCV. core. IPEM.; PF016062; HCV. NSI; IPPC0DOM; PD186062; HCV. NSI; IPPCODOM; PD186062; H
                                                                                                               56476 MW;
                                                                                                                                                                         89.68;
                                                                                                                                                                                                         24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transmembrane;
INIT_MET
                                                 CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVJ5
CARBOHYD
CARBOHYD
                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
SFFF
                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                         a
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation use by mon-profit institutions as long as its content is in no way entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            *** MEDLINE-9223032; PubMed-1314459; A Ckamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., A Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Trull-length sequence of a hepatitis C virus genome having poor signetypes. ", Forology 188:331-341(1992).

***Line of the state of the post of the state of the 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MUG-1992 (Rel. 23, Created)
01-MUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein (Cortains: Capsid protein (Envelope protein M); Major envelope protein (Broshope protein M); Major envelope protein NSI and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus (isolate HC-J7) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                         (POTENTIAL).
                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                               (POTENTIAL).
                                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL)
                                              POTENTIAL)
                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                    81207 MW; 3AF699D82AD501B1 CRC64;
N-LINKED (GLCNAC)
                                                                                                                  (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 138; DB 1;
Pred. No. 2.3e-12;
3; Mismatches 1
                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              737 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPOKRNORNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                           89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 24; Conservative
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 22, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
670-69) (GPN6); Protein F1 (GP32) (GP35); Envelope glycoprotein E2
670-69) (GPN6); Protease/helicase NS3 (P70) (Hepacivirin)
6EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
6EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepacitis C virus (Lasolate BK) (HCV).
                                                                                                                                                                  CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/EZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSZ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; ** Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 737;
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01550; HCV_env; 1.
Probom; PD186062; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Transmembrane; Nonstructural protein;
Transmembrane; Nonstructural protein;
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 1;
Pred. No. 2.3e-12;
3; Mismatches 1.
                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                       N-LINKED
N-L
                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24; Conservative.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                           2333
3053
3053
444
423
7448
534
558
558
558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            carriers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          from human
                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                               TRANSMEM
                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
POLG_HCVBK
                                                                                                                                                                    CHAIN
                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FT FT FT FT SO ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.
MEDLINE-98227846; PubMed-9568891;
Yan Y. Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
virus: a 2.2 - A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998 St.
-- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBLC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
                                                                               X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE=97015088; PubMed=6861916;
Love R.A., Parge H.E., Mickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
The crystal structure of hepatitis C virus NS3 proteinase reveals a
trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR007095; RNA_pol_DS_PS. IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Interpro; IPR002531; HCV_NS1.
Interpro; IPR002518; HCV_NS2.
Interpro; IPR0004109; HCV_NS3.
Interpro; IPR0001495; HCV_NS44.
Interpro; IPR001496; HCV_NS54.
Interpro; IPR002868; HCV_NS58.
Interpro; IPR002166; HCV_NS58.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01543; HCV_caps1d;
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_env;
HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1GX6; 10-APR-02.
1QUV; 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1JXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-NOV-99.
09-APR-02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1A10; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                529.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB
     RANGE RANGE
```

```
RESULT
          SFFFF
                                                                                                               셤
                                                                        REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPPIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-LIRECTED RNA POLYMERASE (POTENTIAL).
                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Halicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                    POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                   N-LINKED (GLCNAC,
                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                             209
234
250
305
305
417
423
4430
448
532
540
556
576
                                                                                                                                                                                                                                                                                                                                      645
2041
2077
2240
                                                                                                                                                                                                                                                                                                                                                                                                                      1074
1076
1081
1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1097
1094
1094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         103
                                                                                                                                                                                                                                                                                                                                                                                     1031
1039
1050
1059
1068
1075
                                                                                                                                                                                                                                                                                                                                                                                                                                              1082
1086
1090
1093
1093
1109
1108
11120
1122
1122
1133
                                                                  3D-structure.
INIT_MET
                                                                                                                                                                TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
STRAND
                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                        CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                            HELIX
STRAND
STRAND
STRAND
TURN
                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                      TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HELIX
TURN
TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                               HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TURN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way entities and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Value 1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
GROWE polyprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3-4-22.-); Proteass/helicase NS3 (P70) (Hepacivirin)
(EC 3-4-21.98); Nonstructural protein NSA (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NSA (P56); Nonstructural protein
NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISCUSSION OF SEQUENCE.

MEDLINE-91192160; PubMed-1849488;

MEDLINE-91192160; PubMed-1849488;

MATCO N. Hillkata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

Ohkoshi S., Shimotohno K.;

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese Polyprotecular o
                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                {RNA}(N).
-!-SUBURT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                           Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                         327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                          89.6%; Score 138; DB 1;
llarity 85.7%; Pred. No. 1.1e-11;
Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                 1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D90208; BAA14233.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
     1188
1197
1202
1204
                                                                                                                  1688
                                                                                                                                            3010 AA;
                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A39253; GNWVC.
     1187
1189
1198
1203
1680
                                                                                                                                                                                                / Mu.
Local Sim-
24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVJA
ID POLG_HCVJA
AC P26662;
                                                                                                          STRAND
                                                                                                                                                                                             Query Match
TURN
                                                      HELIX
                                                                                      TURN
```

```
1 PKPOKRNORNINRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Prince A.M.;
                                                                                                                                                                 POLG_HCVH
P27958:
                                                                                                                       RESULT 9
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE,HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 138; DB 1; Length 3010; Pred. No. 1.1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . .) (POT MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED GECNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                        rPro; IPR007095; RNA_pol_DS_PS.
rPro; IPR007094; RNA_pol_PSvir.
; PF01543; HCV_capsid; 1.
; PF01542; HCV_core; 1.
; PF01550; HCV_env; 1.
                                                                                                                     | IPR002519 | HCV_env.
| IPR002519 | HCV_NS1.
| IPR002518 | HCV_NS2.
| IPR004109 | HCV_NS3.
| IPR00145 | HCV_NS4a.
| IPR001490 | HCV_NS4b.
| IPR00166 | HCV_NS6a.
| IPR002166 | HCV_NS6a.
                                                                                HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00271; helfcase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4a;
HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC;
                                                                            InterPro; IPR002522;
InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2077 207
2240 224
2788 278
3010 AA;
                                                           IPR001410;
                 MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF02907;
Pfam; PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01001;
                                                                                                                                                                                                                           InterPro;
InterPro;
                                                               InterPro;
                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                            InterPro;
                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
VP_BIND
3ITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
ODRA NO DE LA COLLA DELLA DELL
```

Gaps

ö

1; Indels

3; Mismatches

85.78;

24; Conservative

Best Local Similarity Matches 24; Conserv

```
e are no restrictions on its
s its content is in no way
Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    between the Swiss Institute of Bioinformatics and the FWBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no warmodified and this statement is not removed. Usage by and for commercial
                                                                                                                                     15.52P-203 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15.52P-203 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99.-); Protease/helicase NS3 (P70) (Hepacivirin)
NS4 (P70); Nonstructural protein NS4A (P4); Nonstructural protein
NS5B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-99154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N)
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
--- PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates.";
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                           PRT; 3011 AA
SEQUENCE FROM N.A. MEDLINE-92052256; PubMed-1658800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Structure 6:89-100(1998).
                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11108;
```

```
Gaps
                                                · ·) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                       N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     llarity 85.7%; Pred. No. 1.1e-11; Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1627 16
1635 16
1640 16
3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 24; Conserv.
                             CARBOHYD
CARBOHYD
STRAND
TURN
HELIX
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
HELIX
HELIX
HELIX
TURN
STRAND
HELIX
TURN
STRAND
TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                          TURN
TURN
STRAND
STRAND
STRAND
TURN
STRAND
HELIX
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
                               ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
                                          DR EMBL; M67463; AAA45534.1;

DR PRE A36614; GNWCH.

DR PDB; 1ARV, 16-FEB-99

DR PDB; 1ARV, 16-FEB-99

DR PDB; 1ARV, 17-UN-98

DR PDB; 1ARV, 17-UN-98

DR PDB; 1ARV, 17-UN-98

DR PDB; 1ARV, 17-UN-98

DR PREOPS; 829.001;

DR INTEPTO; PRRO02521; HCV_CAPSIG

DR INTEPTO; PRR002521; HCV_CAPSIG

DR INTEPTO; PRR002519; HCV_CAPSIG

DR INTEPTO; PRR002519; HCV_MS1.

DR INTEPTO; PRR002519; HCV_MS2.

DR INTEPTO; PRR002519; HCV_MS3.

DR INTEPTO; PRR002569; HCV_MS3.

DR INTEPTO; PRR00269; HCV_MS3.

DR INTEPTO; PRR00269; HCV_MS3.

DR INTEPTO; PRR00269; HCV_MS3.

DR INTEPTO; PRR00269; HCV_MS3.

DR Pfam; PPG1543; HCV_COPF; I.

DR Pfam; PPG1543; HCV_COPF; I.

DR Pfam; PPG1559; HCV_MS3.

DR Pfam; PPG1560; HCV_MS3.

DR Pfam; PPG1059; HCV_MS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5B.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11711
1972
3011
3692
11083
11083
11165
1237
2096
2096
417
423
4413
4413
4413
4413
532
532
555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
```

\$ 3033

STANDARD;

```
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
DORR

SPECTATION SPECTATION SECTION SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-1., Suglyama Y., Kurai K., Lizuka H.,
A Vachida A., Miyakawa Y., Mayumi M.;
A vachida A., Miyakawa Y., Mayumi M.;
A vachida A., Miyakawa Y., Mayumi M.;
A uclectide sequence of the genomic RNA of hepatitis C virus isolated
in of the genome of the genomic RNA of hepatitis C virus isolated
in of the genome of the genomic RNA of hepatitis C virus isolated
in of the genome of the genomic RNA of the genome of t
                        01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Proteasse/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21-98); Nonstructural protein NS5A (P4); Nonstructural protein NS5A (P65); Nonstructural protein NS5A (P66); Nonstructural protein NS5A (P670); Nonstructural 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1- SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BY AND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPROUZEGO, HCYMARP
IPROUZEGO, HCYMARP
IPROUZEGO, HGLICASE, C.
IPROUZEGO, RNA_DOL_DS_PS.
IPROUZEGO, RNA_DOL_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SSP; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01006;
PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01538;
PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein (P22);
Envelope glycoprotein E1 (GP22) (GP55); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P57); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J8) (HCV).
                                                                                                                                                CELLULAR AMINOPEDITIDASE.
CELLULAR AMINOPEDITIDASE.
CELLULAR AMINOPEDITIDASE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00298; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                           Coré protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1
                                                                                                                                                                                                                                                                                                                                                               SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 1;
Pred. No. 1.1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                         DECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2038
                                                                                                                                                                                                                                                        1010
1619
1866
2017
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2038 20;
2811 28;
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                              347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVJ8
P26661;
```

```
42. ESD
```

```
10.4PR-1993 (Rel. 25, Created)
11.5Ep-2003 (Rel. 25, Last sequence update)
12.5Ep-2003 (Rel. 42, Last annotation update)
13.5Ep-2003 (Rel. 42, Last annotation update)
15.5Ep-2003 (Rel. 42, Last annotation update)
15.5Ep-2003 (Rel. 42, Last annotation update)
16.5Ep-2003 (Rel. 42, Last annotation update)
17.5Ep-2003 (Rel. 42, Last annotation update)
18.5Ep-2003 (Rel. 42, Last annotation u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.6%; Score 138; DB 1; Length 3033;
85.7%; Pred. No. 1.1e-11;
tive 3; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED !
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MW; 1A173E7E3381FD1A CRC64;
                                                                                                                         CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                            N-LINKED (GLCNAC
                                                                                                                                                                                        ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                 N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                            ACT_SITE
ACT_SITE
                                                                                                                                                 ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                        RANSMEM
                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_HCVTW
                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local S
Matches 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
POLG_HCVTW
                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration the European Bioinformatics of Bioinformatics and the EMBL outstation use by non-profit institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
      SEQUENCE FROM N.A.
MEDINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Full-length Sequence of a hepatitis C virus genome having poor
genotypes...
                                                                                                                                          Virology 188131-341(1992).

-!- FUNCTION: THE SMALL PROFEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the Postion, Cys or Thr in P1 and Ser or Ala in P1'.
                                                                                                                                                                                                                                                                                                                                     SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CABSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
PROTESSE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Cat protein; Envelope Protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Interpro: IPR007095; NN_DOL_DS PS Pfam; PF01543; HCV_caps1d; IPF007094; RNA_DOL_DS_PS Pfam; PF01542; HCV_caps1d; IPFam; PF01542; HCV_caps; IPFam; PF01539; HCV_env; IPFam; PF01560; HCV_NS1; IPFam; PF01006; HCV_NS2; IPFam; PF01006; HCV_NS4; IPFam; PF01006; HCV_NS4; IPFam; PF01001; HCV_NS4b; IPFam; PF01001; HCV_NS4b; IPFam; PF01001; HCV_NS4b; IPFam; PF010099; VITAL_RGRP; IPFAM; SMART; SM00499; VITAL_RGRP; IPFODOM; PD186062; HCV_NS1; IPFODOM; PD18606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Interpro; IPR002522; HCV_capsid.
Interpro; IPR00251; HCV_capsid.
Interpro; IPR002519; HCV_NSI.
Interpro; IPR002519; HCV_NSI.
Interpro; IPR0040518; HCV_NSI.
Interpro; IPR004109; HCV_NS2.
Interpro; IPR004409; HCV_NS3.
Interpro; IPR004409; HCV_NS4.
Interpro; IPR002668; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10988; BAA01761.1; -. PIR; A40250; GWWJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
```

BY A

Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein (Contains Capsid protein C (Core protein) (P22);
67-68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); Nonstructural protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-92295714; PubMed-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijkata M., Ishimura Y., Shimotchno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VILLUS RES. 23:39-53 (1992).

-1 FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4B AND NS4B ARE HYDROPHOBLE. SUGGESTING A POSSIBLE MEMBRANE-RELFED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1 CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or 31u in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-1 CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -1- SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                          (POTENTIAL)
                                                  POTENTIAL)
                                                                            POTENTIAL)
                                                                                             (POTENTIAL)
                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                 Score 131; DB 1; Length 3010;
Pred. No. 1.2e-10;
3; Mismatches 1; Indels
                                                                                                                                                                                                                              N-LINKED (GLCNAC. ..) (POTE N-LINKED (GLCNAC. ..) (POTE N-LINKED (GLCNAC. ..) (POTE MW; AAD267055CDFE215 CRC64;
                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                 85.1%;
                                                                                                                                                                                                                                                                                                      327047
                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 85....
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                     623
645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
NCBI_TaxID=31642;
                                                                                                                  576
623
645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJT
000269:
                                                                     CARBOHYD
CARBOHYD
                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                  CARBOHYD
                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01538; HCV_NS2; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdFp; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA properties.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410; DEAD.
InterPro; IPR001213; HCV-capsid.
InterPro; IPR001213; HCV-capsid.
InterPro; IPR001213; HCV-care.
InterPro; IPR001213; HCV-care.
InterPro; IPR001219; HCV-care.
InterPro; IPR001219; HCV-care.
InterPro; IPR001409; HCV-capsid; I-FSVIr.
                                                                                                                                                                                                                                                                           EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1083
1107
1165
11237
1319
196
209
                                                                                                                                                                                                                                                                                                                    PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                 GNWVTW
                                                                                                                                                                                                                                                                                                                                                               MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01542; H
Pfam; PF01539; H
Pfam; PF01560; H
Pfam; PF01538; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1107
1165
1230
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              196
209
233
234
250
305
417
                                                                                                                                                                                                                                                                                                   PIR; A40244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
:NIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

ø

2002

SEQUENCE

g

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJUN PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN IS (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/FELLCASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR007095; RNA_pol_DS_PS
InterPro; IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                         HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                         InterPro; IPR002518; HCV_NS2
InterPro; IPR004109; HCV_NS3
InterPro; IPR000145; HCV_NS4
InterPro; IPR001409; HCV_NS5a
InterPro; IPR002868; HCV_NS5a
InterPro; IPR002166; HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam: PF01543; HCV_capsid; 1. Pfam: PF01543; HCV_capsid; 1. Pfam: PF01542; HCV_capsid; 1. Pfam: PF01554; HCV_capsid; 1. Pfam: PF01556; HCV_NS; 1. Pfam: PF01506; HCV_NS; 1. Pfam: PF01006; HCV_NS3; 1. Pfam: PF01001; HCV_NS4; 1. Pfam: PF01001; HCV_NS4; 1. Pfam: PF01506; HCV_NS5a; 1. Pfam: PF001506; HCV_NS5a; 1. Pfam: PF00998; viral_RARP; 1. PCDOm: PD186062; HCV_NS1; 1. PCD186062; 
                                                                                                                                                                                                IPR002522; H
IPR002521; H
IPR002519; H
IPR002531; H
                                                         PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2041
2077
22240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1165
1230
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure.
INIT_MET
                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
STRAIN-New Zealand; TISSUE-Brain, and Skeletal muscle;
MEDLINE-96161994; Pubmed-8776224;
CSOTIOS C., Zolnierowicz S., Bako E., Durbin S.D., Depaoli-Roach A.A.;
Fligh complexity in the expression of the B' subunit of protein phosphatase 2A0. Evidence for the existence of at least seven novel isoforms."
J. Biol. Chem. 271:2578-2588(1996).
I. FUNCITON: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa CONTANT REGULATORY SUBUNIT (SUBUNIT A), THAT ASSOCIATES WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE HITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE B (THE RZ/B/FR55/B55, R3/B''/PR72/PR130/PR59 AND R5/B'/B56 AND R5/B'/B56 AND CELL SIGNALING MOLECULES.
-! SUBCELLUAR LOCATION: Nuclear (By similarity).
-! TISSUE SPECIFICITY: Highly expressed in brain.
-! TISSUE SPECIFICITY: Highly expressed in brain.
                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                          2A5D_RABIT STANDARD; PRT; 586 AA.

Q28653; Q28655;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2011 (Rel. 40, Last annotation update)
56-tine/threonine protein phosphataes 2A, 56 kDa regulatory subunit,
B5 Gelta isoform (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, PK61 delta isoform) (PP2A, B subunit, PK61 delta isoform) (PP2A, B subunit, PK61 delta isoform) (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, B'-qamma).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SH3 BINDING, CLASS I (POTENTIAL), NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Multigene family; Nuclear protein; Repeat.

DOMAIN 21 36 8 X 2 AA APPROXIMATE TANDEM REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Oryctolagus cuniculus (Rabbit).
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1. CAULON: NOMENCLATURE USED IN REF.1 REFERS TO PP2A B SUBUNIT B' GAMMA ISOFORM, WHICH IS CITED AS PP2A B SUBUNIT DELTA-PR61 ISOFORM IN LATER PUBLICATIONS.
                                                                                      Length 3010;
                                                                     Score 130; DB 1; Length 30
Pred. No. 1.6e-10;
3; Mismatches 2; Indels
3010 AA; 326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                      84.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U38193; AAC48532.1; --
EMBL; U38195; AAC48534.1; --
InterPro; IPR002554; B56.
Pfam; PF01603; B56; 1.
                                                                                         Similarity 82.1
23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   507
                                                               Query Match
                                                                                            Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                2A5D_RABIT
                                                                                                                                                                                                                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                          THE TENT TO BE SEED FOR THE TENT TO SEE SEED FOR THE TENT TO SEE SEED FOR THE TENT TO SEE SEE SEED FOR THE TENT TO SEED FOR THE TENT TO SEE SEED FOR THE TENT TO SEED FOR THE TENT TO SEE SEED FOR THE TENT TO S
                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
```

```
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE-Colon, Eye, and Kidney.

X MEDLINE-22388257; PubMed=12477932;

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

A Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

A Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

A Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

A Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

A Diatchenko L., Morares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

B Diatchenko L., Woldin T.B., Toshlyuki S., Carninci P., Prange C.,

R Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

R Raba S.S., Worley K.C., Hale S., Garcia A.M., Gabbs R.A.,

A Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

A Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Halting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

R Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

R Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Fetal brain;
MEDLINE-96355607; PubMed-8703017;
MCCright B., Rivers A.M., Audlin S., Virshup D.M.;
"The B56 family of protein phosphatase 2A (PP2A) regulatory subunits encodes differentiation-induced phosphoproteins that target PP2A to
                                                           Gaps
                                                                                                                                                                                                014738; 000494; 000696; 015171; 000.
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2003 (Rel. 42, Last annotation update)
Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PPAA, B subunit, B' delta isoform) (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, R5 delta isoform) (PP2A, B subunit, B56 delta isoform) (PP2A, B subunit, B61 delta isoform)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    550-559; 573-580 AND 584-601 (DELTA-1).
TISSUE-Bone marrow, and Brain cortex;
MEDLINE-96159032; PubMed-856219;
Tanabe O., Nagase T., Murakani T., Nozaki H., Usui H., Nishito Y.,
Hayashi H., Kagamiyama H., Takeda M.;
"Molecular cloning of a 74-kDa regulatory subunit (B'' or delta) of human protein phosphatase 2A.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE-Brain cortex;
MEDLINE-97324098; PubMed-9180267;
Manabe O., Gomez G.A., Nishito Y., Usui H., Takeda M.;
Tanabe O., Gomez G.A., Nishito Y., Usui H., Takeda M.;
"Molecular heterogeneity of the cDNA encoding a 74-kba regulatory
subunit (B' or delta) of human protein phosphatase 2A.";
FEBS Lett. 408:52-56(1997).
                                                                                                                                                                                                                                                                                                                                          Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                         11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORM DELTA-2), AND SEQUENCE OF 501-508;
                             Length 586;
                                                         Indels
68090 MW; E149A309CDDA7495 CRC64;
                                                                                                            .29
                                                                                     1 PKPQKRNQ-RNTNRRPQD-----VKFPGGGQIV 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-2)
                                                         .
.
                                                                                                      1;
                             DB
                                                                                                                                                                                            602 AA
                       Score 54.5; Di
Pred. No. 2.7;
                                                       9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (ISOFORM DELTA-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            both nucleus and cytoplasm.";
J. Biol. Chem. 271:22081-22089(1996)
                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EBS Lett. 379:107-111(1996).
                             35.4%;
34.2%;
                                                           13; Conservative
                                                                                                                                                                                            STANDARD;
586 AA;
                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=9606;
                                                                                                                                                                                          2A5D_HUMAN
                                                                                                                  28
SECUENCE
                             Query Match
                                                           Matches
g
                                                                                                                                                                                            g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; $68686; $68686.

Genew; HGNC:9312; PPP2R5D.

GK; O14738; -.

MIM; 601646; -.

GO; GO:0005634; C:nucleus; TAS.

GO; GO:0007399; P:neurogenesis; TAS.

InterPro; IPR002554; B56.

Pfam; PF01603; B56; I.
                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
MASSING (In 190form Delta-3).
/FTIGH-VSP_005110.
Missing (in isoform Delta-2).
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: THE B REGULATORY SUBMYIT MOGHARE SUBSTRATE SELECTIVITY AND CALATYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC SIZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISOId=Q14738-3; Sequence=VSP_005110;
TISSUE SPECIFICITY: ISOFORM DELTA-2 IS WIDELY EXPRESSED. ISOFORM
DELTA-1 IS HIGHLY EXPRESSED IN BRAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8 X 2 AA APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                                  SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein; Phosphorylation; Alternative splicing; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASS I (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- INDUCTION: BY RETINOIC ACID; IN NEUROBLASTOMA CELL -1- PTM: AT LEAST ISOFORM DELTA-1 IS PHOSPHORYLATED ON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           F15F71AF4E565387 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Event-Alternative splicing; Named isoforms-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   005111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IsoId=Q14738-2; Sequence=VSP_005111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SH3 BINDING,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IsoId-014738-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FTIG-VSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL: BC001095; AAH01095.1; -. EMBL: BC001175; AAH01175.1; -. EMBL: BC010692; AAH10692.1; -. PIR; S68686; S68686.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AB000634; BAA20381.1; -. EMBL; AB000635; BAA20382.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69991 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; L76702; AAB69751.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D78360; BAA11372.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      530
565
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Name-Delta-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Name-Delta-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Name-Delta-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         85
                                                                                                                                                                                                          COMPARTMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -!- PTM: AT I
RESIDUES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Multigene
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
```

DB 1; Length 602;

Score 54.5; DE Pred. No. 2.7;

35.4%;

us-09-491-146a-29.rsp

Matches 13; Conservative 9; Mismatches 5; Indels 11; Gaps ò

Ωp

Search completed: August 7, 2003, 11:20:02 Job time : 5.90909 secs

hepatitis hepatitis

hepatitis hepatitis

hepatitis

081yr6 068712 068712 068708 068679 068683 068683 068684 068684 068685 081ys0 081ys0 081ys0 081ys0 081ys0 081ys0 081ys0

hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis

084p52 084p88 094t56 084p90 084p64 084p64 084p55

hepatitis hepatitis

hepatitis

09qt58

hepatitis hepatitis

hepatitis

hepatitis hepatitis hepatitis

hepatitis hepatitis

hepatitis

hepatitis

hepatitis

08jyr5 08jyr7

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              יו-האויויין (TrEMBLrel. 03, Last sequence update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
Core, env.and part of E2/NS1 (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           415 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                       068712
068682
068682
068679
068687
068684
068687
068685
068707
068685
068707
                                                                                                                                                                                                                                                      080P70
080P54
090T57
                                                                                                                                                                                                                                                                                                                             090T56
080P90
080P64
080P55
                                                                                                                                                                                                                                                                                                     Q8QP52
Q8QP88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-1997 (TrEMBLrel. 03, 01-MAY-1997 (TrEMBLrel. 03,
 PRELIMINARY;
 [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-VN235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Okamoto H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P87751
P87751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1

      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1
      1

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P87751 hepatitis c 092530 hepatitis c 080772 hepatitis c 068137 hepatitis c 096983 hepatitis c 099199 hepatitis c 0807w0 hepatitis c 0807w0 hepatitis c 0807w8 hepatitis c 0807w8 hepatitis c 0807w8 hepatitis c 0807w8 hepatitis c 080817 hepatitis c 068310 hepatitis c 068306 hepatitis c 068306 hepatitis c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hepatitis
hepatitis
                                                                                                                    (without alignments)
281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                      Search time 25.6364 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28jyr9
                                                                                                                                                                                                                                                                                                   830525
            GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                 1 number of hits satisfying chosen parameters:
                                                                                                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                    830525 segs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                      August 7, 2003, 11:05:41;
                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                        protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9E983
Q9J3H8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0901Y9
08V7V3
08V7W0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08V7W8
08V7W5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q68311
Q68306
Q68309
Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q8JWL7
                                                                                                                                                                                                                                      Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_organelle:*sp_phage:*
                                                                                                                                                                                                                                                                                                                               seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_archea:*
sp_bacteria:*
                                                                                                                                                               US-09-491-146A-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_fund1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                     SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   103
3010
3010
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
191
191
                                                                                                                                                                                                                            BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0

1000.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 141
140
140
140
1138
1138
1138
                                                                                                                                                               Title:
Perfect score:
                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       146
142
141
141
                                                                        OM protein
                                                                                                                                                                                                                                                                                                                               Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                      Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ٠.
و
```

2

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                  lesions.
                                                                                                                                                                                                                          Q8V7Y2;
                                                                                                                                                                                                               Q8V7Y2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              068137
                                                                                                           Matches
                                                                                                                                                                                        RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
                                                                                                  Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   068137
          NA WAY SO
                                                                                                                                                     a
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                             Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F., Miyakawa Y., Mayumi M.,
                                                                                   Gaps
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
NON_TER 415 415 415 SEQUENCE 415 AA; 45301 MW; 39B2040A86517D1R CRC64.
                                                                                                                                                                                                                                                                                                                                                                                                                The entire nucleotide sequences of three hepatitis C virus isolates in genetic groups 7-9 and comparison with those in the other eight greatic groups.;
J. Gen. VITOL. 79:1847-1857(1998).
I. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS.
PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MINA (BY SIMILARITY).
                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                 ô
                                                            Length 415;
                                                                                 Indels
                                                                                                                                                                                                                                                                                                                          Okamoto H.;
Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
                                45301 MW; 39B2040A86517D1B CRC64;
                                                                                                                                                                                                  01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                         ; Score 154; DB 12;
; Pred. No. 1.7e-14;
0; Mismatches 0;
                                                                                                                                                                                 PRT; 3013 AA.
                                                                                                 1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                            InterPro; IPR002513; HCV_NS1.

R InterPro; IPR004109; HCV_NS3.

R InterPro; IPR004109; HCV_NS3.

R InterPro; IPR004109; HCV_NS3.

InterPro; IPR001490; HCV_NS4a.

InterPro; IPR002868; HCV_NS5a.

InterPro; IPR002166; HCV_NS5a.

InterPro; IPR007165; RNA_DOl_DS_PS.

InterPro; IPR007095; RNA_DOl_DS_PS.

Pfam; PP01543; HCV_Corpeil

Pfam; PP01542; HCV_Corpeil
                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-98378034; PubMed-9714232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_capsid.
HCV_core.
HCV_env.
                                                       Query Match
Best Local Similarity 100.0%;
Matches 28; Conservative 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D84263; BAA32665.1;
HSSP; P27958; 1A1V.
                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519;
                                                                                                                                                                                                                                    Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00998; Vir
ProDom; PD186062;
                                                                                                                                                                                                                                                                                                                 STRAIN-VN235;
                                                                                                                                                                                                                                                                                                                                                                        STRAIN-VN235
                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                             092530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
Pfam;
                                                                                                                                                        RESULT 2
092530
                                                                                                  ò
                                                                                                                     g
                                                                                                                                                                              44446
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
SWART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
COAT protein; Envelope protein; 01ycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3013 AA; 328196 WW; C9EE9C0231E86EAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                    Hepptitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                         ó
                                                                                                        100.0%; Score 154; DB 12; Length 3013; 100.0%; Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kato N.; "Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 191;
                                                                                                                                        0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191 191
191 AA; 20703 MW; 3215979FE3F45EB7 CRC64;
                                                                                                                                                                                                                                                                                             01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94.8%; Score 146; DB 12; 92.9%; Pred. No. 1.2e-13;
                                                                                                                                                                                                                                                                                                                                        Core protein (Genome polyprotein) (Fragment).
                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                5 PKPOKRNQRNINRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein (Fragment).
                                                                                                                 Local Similarity 100.
                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. STRAIN-SA6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON_TER 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Simi
Matches 26;
```

103 AA; 11641 MW; B59832ACDA5085EA CRC64;

us-09-491-146a-29.rspt

```
SEQUENCE
                                                                                                                           093348;
                                                                                                                 0933н8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                 RESULT 6
Q9J3H8
SO
                                                                     a
                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                   ö
                                                                   MEDLINE-92279243; PubMed-1317578;
Bukh J., Purcell R.H., Miller R.H.;
'Sequence analysis of the 7, noncoding region of hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 89:4442-4946(1992).
                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          K
                                                                                                                                MEDLINE-93376778; PubMed-8396266;
Bukh J., Purcell R.H., Miller R.H.;
*At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative El gene of isolates collected worldwide.";
Proc. Natl. Acad. Sci. U.S.A. 90.824-8238(1993).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY MILLARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                 Score 142; DB 12; Length 191;
Pred. No. 4.5e-13;
2; Mismatches 1; Indels C
    Bukh J., Purcell R.H., Miller R.H.; "Sequence analysis of the core gene of 14 hepatitis C virus genotypes.";
                                                                                                                                                                                                                                                                       191
20682 MW; OEB8000CF1F902EB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                 01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update).
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                  Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
                                                                                                                                                                                                                                                                                                                                                                                                  103 AA.
                                                                                                                                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTQRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
MEDLINE-94336721; PubMed-8058787;
                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein (Fragment).
                                                                                                                                                                                                                                                                                                 92.2%;
                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 89.39
                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        103
                                                                                                                                                                                                                                                              Polyprotein.
NON_TER 191 1
SEQUENCE 191 AA;
                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                 SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HC-N3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
                                                                                                                           STRAIN-SA6;
                                                              STRAIN-SA6
                                                                                                                                                                                                                                                                                                                                                                                                          Q9E983;
                                                                                                                                                                                                                                                                                                                                                                                                 Q9E983
                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
Q9E983
STERRERES
                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                   1255
1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         progression.";
Submitted (Nov-1999) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN B. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, AF207755; AAF65945.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C. "Characteristics of hepatitis C viral genome associated with disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00190, CYTCCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Rn-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 326421 MW; B109B6487CD206E8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                 ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 3010;
                          Length 103;
                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                            01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                     Score 141; DB 12;
Pred. No. 3.3e-13;
2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.6%; Score 141; DB 12;
89.3%; Pred. No. 1.2e-11;
tive 2; Mismatches 1;
                                                                                                                                                                                                                                                         PRT; 3010 AA
                                                                                                    6; HCV_RdRP.
5; RNA_pol_DS_PS.
4; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000345; CytC_heme_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid.
                     91.6%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR007094; RNA_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a;
                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SM00487; DEXDC;
Query Match
Best Local Similarity
Local 25; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR001410;
                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4SSP; P26663; lJXP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-MD14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nterPro:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nterPro
```

PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28 

~1

g

```
Nagayama K., Kurosaki M., Enomoto N., Maekawa S., Miyasaka Y., Sakamoto N., Fukuma T., Tazawa J., Izuui N., Marumo F., Sato C.; Submitted (JuL.1999) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                     Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
Tazawa J.i., Izumi N., Marumo F., Sato C.;
"Time-related changes in full-length hepatitis C virus and hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS50507, RDRP POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.6%; Score 141; DB 12; Length 3010;
89.3%; Pred. No. 1.2e-11;
1.ve 2; Mismatches 1; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             D816D3BBBF14EE46 CRC64;
                                                                                                                                                          01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus.
                                                                                                                           PRT; 3010 AA
PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                 PKPQRRTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL' AF165046; AND MANA (BY SIMILARITY)
HSSP; P2663; 1JXP.
INTERPRO; IPR001410; DEAD.
INTERPRO; IPR002521; HCV_capsid.
INTERPRO; IPR002521; HCV_capsid.
INTERPRO; IPR002521; HCV_capsid.
INTERPRO; IPR002511; HCV_NS1.
INTERPRO; IPR002511; HCV_NS2.
INTERPRO; IPR002101; HCV_NS2.
INTERPRO; IPR001409; HCV_NS4.
INTERPRO; IPR001409; HCV_NS4.
INTERPRO; IPR001409; HCV_NS4.
INTERPRO; IPR00166; HCV_NS4.
INTERPRO; IPR00166; HCV_NS4.
INTERPRO; IPR007094; RNA_POI_DS_PS.
INTERPRO; IPR007094; RNA_POI_DS_PS.
INTERPRO; IPR007094; RNA_POI_PSVIT.
Pfam; PF01543; HCV_Capsid; I.
Pfam; PF01560; HCV_NS1; I.
Pfam; PF01006; HCV_NS4; I.
Pfam; PF01006; HCV_NS4; I.
Pfam; PF01001; HCV_NS4; I.
Pfam; PF01006; HCV_NS4; I.
                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-20013325; PubMed-10544098;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3010 AA; 326366 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virology 263:244-253(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 89.3
Matches 25, Conservative
                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SM00487; DEXDC:
                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-MD1-2;
                                                                                                                                                                                                                                                                                                                                                STRAIN-MD1-2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                           0901Y9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART;
                                                                                        RESULT 7
à
                                  q
                                                                                                                             DR ADD BR DR KKW KKW SQ SS SQ SS
```

```
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062026; BAB83365.1; -.

InterPro; IPR00252; HCV_capsid.

InterPro; IPR00251; HCV_core.

Pfam; PF0134; HCV_capsid.

Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ø
                                      0807v3;
01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
01-007-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses: ssrNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (MAX-2001) to the EMBL/GenBank/DDBJ databases.
-1- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AB062019; BABB3358 1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
PinterPro; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                               Kato N.;
"Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Hepatitis C virus quasispecies in cancerous and non-cancerous lesions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90.9%; Score 140; DB 12; Length 191; 89.3%; Pred. No. 8.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 191 191 191 191 AA; 20700 MW; DDEF95E878228E07 CRC64;
                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                              191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQKKTNRNTNRRPQDVKFPGGGQIVG 32
                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 89.3 es 25; Conservative
                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                SEQUENCE FROM N.A. TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   191
                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                               lesions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8V7W0
Q8V7W0;
                           08V7V3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0877W0
                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

ö

Gaps

ö

S

ö

```
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ogata S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QBJWL7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8JWL7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             068311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q8JWL7
   SO TEN DOR CCC CCC ST WE DOR SO ST WE DOR SO ST WE SO ST WE ST WE SO ST WE WE ST WE WE ST WE WE ST WE WE ST WE ST WE ST WE ST WE WE ST WE WE ST WE ST WE ST WE ST 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (MAX-2001) to the EMBL/GenBank/DDBJ databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062011; BAB83350.1; -..

InterPro: PRR00252; HCV_capaid.

InterPro: IPR002521; HCV_core.

Pfam; PF01343; HCV_core.

Pfam; PF01342; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Hepatitis C virus quasispecies in cancerous and non-cancerous lesions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                             Score 140; DB 12; Length 191;
Pred. No. 8.9e-13;
1; Mismatches 2; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 140; DB 12; Length 191;
Pred, No. 8.9e-13;
1; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20667 MW; IEEF95EA0E115BFA CRC64;
                          POLYProtein.
NON_TER 191 191
SEQUENCE 191 AA; 20690 MW; DDEF9785E8228E07 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQKKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                          1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90.9%;
                                                                                                                                                          90.9%;
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                        Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 89.3 nes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rissue-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8V7W5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8V7W8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08V7W5
                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
Q8V7W8
ID Q8V7W
AC Q8V7W
DT 01-MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8V7W5
Q8V7W5
Q8 DAC
Q9 DAC
Q9 DDT
Q1D DT
Q1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                             ò
```

```
"Comparative Sequence Analysis of the Core Protein and Its Frameshift Product, the Frotein, of Hepatitis C Virus Subtype 1D Strains Product, the Frotein, of Hepatitis C Virus Subtype 1D Strains obtained from Patients with and without Hepatocellular Carcinoma."; Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment).
Hepatitis C virus type 1b.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN A AND MANA (BY SIMILARITY).

EMBL: AB062014: BAB83353.1; -
InterPor: IPR00252; HCV_capsid.
InterPor: IPR00252; HCV_capsid.
InterPor: IPR00252; HCV_core.
Pfam: PF01543; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                             Score 140; DB 12; Length 191;
Pred. No. 8.9e-13;
1; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90.3%; Score 139; DB 12; Length 191; 85.7%; Pred. No. 1.3e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 191 20804 MW; F5CDCB21037433C8 CRC64;
                                                                                                                                                                                                                                                                                     191 191 20690 MW; DDEF9785E8228E07 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (TrEMBLrel. 22, Last sequence update)
(TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PKPQKKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AB077728; BAC02463.1; -...
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                             90.9%;
llarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel. 01, (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. STRAIN-L27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
NCBI_TaxID=31647;
                                                                                                                                                                                                                                                             Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q68311
Q68311;
Q1-NOV-1996 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996
01-OCT-2002
```

PRT;

PRELIMINARY;

```
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Heparitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN-HCV-BB10;
                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-HCV-BB5; Songalvilal S., Kanistanon D., Kunkitti R.; Identification and characterisation of Thal isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIVCOPROFEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; U23744; AAA65011.;
                                                                                                                                                                                                                                                 Virus.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!-SUBGUIT: THE VIRLON OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY ILIPOPROTEIN THE NVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; 023749; AA65056.1;
-InterPro; IPR00252; HCV_caps1d.
Pfam: PF01543; HCV_caps1d.
                                                                                                                                                                                                    Songsivilal S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 12; Length 45; Pred. No. 3.7e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match 89.6%; Score 138; DB 12; Length 45; Best Local Similarity 85.7%; Pred. No. 3.7e-13; Matches 24; Conservative 3; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 45
45 AA; 5015 MW; CC527167096AAAB1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 45
45 AA; 5030 MW; CC4C21ED236AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPOKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              vuery Match
Best Local Similarity 85.77
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                               STRAIN-HCV-BB37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
068306
DD 068306;
DT 01-NOV-
DT 01-NOV-
DD 01
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SO TE BRANES
```

RESULT 15 Q68309

```
ö
                                    VICUS.";
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- SUBMIT: THE VIEND OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                   Score 138; DB 12; Length 46;
Pred. No. 3.8e-13;
3; Mismatches 1; Indels
                                                                                                                                                                                                                                                      46 46
46 AA: 5129 MW; ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Search completed: August 7, 2003, 11:19:01 Job time : 25.6364 secs
                                                                                                                                                                EMBL; U23747; AAA65054.;
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                         ilarity 85.7%;
Conservative
                                                                                                                                                                                                                                                                                                                     89.68;
                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                            Polyprotein.
NON_TER
SEQUENCE 46
                                                                                                                                                                                                                                                                                                                     Local Simple 24; C
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   internal
MOLECULE TYPE: protein
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FRAGMENT TYPE:
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; ORGANISM:
US-08-921-887-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-921-887-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ANTI-SENSE:
Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 29, App.
                                                                                                                                          August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sednence sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                    328717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
/cgn2_6/ptodata/1/laa/5B_COMB.pep:*
/cgn2_6/ptodata/1/laa/6A_COMB.pep:*
/cgn2_6/ptodata/1/laa/6B_COMB.pep:*
/cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
                  5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-921-887-29
US-07-681-703B-26
US-08-485-500-26
PCT-US91-02370-26
PCT-US92-07813-1
US-07-681-703B-20
US-08-405-500-20
PCT-US91-02370-20
PCT-US91-02370-20
US-08-405-500-20
PCT-US91-02370-20
US-08-445-500-18
PCT-US91-02370-18
US-08-485-500-18
US-08-485-500-18
US-08-485-500-18
US-08-485-500-18
US-08-485-500-18
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
US-08-444-818-103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US91-02370-16
                                                                                                                                                                                                                                                                                                                                                                                                                               ital number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-07-681-703B-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -US91-02370-12
                                                                                                                                                                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                         328717 segs, 42310858 residues
                  GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                      protein search, using sw model
                                                                                                                                                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length: 0
length: 2000000000
                                                                                                                                                                                                                                   US-09-491-146A-29
154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sed
                                                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB :
Maximum DB :
                                                                                                      OM protein
                                                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                                                               Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
```

```
Sequence 29, Application US/08921887
Patent No. 603071
GENERAL INFORMATION:
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAM
WUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES 6 ASKEW, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                               Appl
Appl
Appli
                                                  Sequence 7, Agenence 12, Sequence 12, Sequence 12, Sequence 13, Sequence 36, Sequence 36, Sequence 10,
                                                                                                                                       Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
                                 Sequence
      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 154; DB 3;
Pred. No. 4.9e-15;
US-08-485-500-12
US-08-485-500-14
US-08-485-500-14
US-07-910-760-12
US-08-440-519-12
US-08-440-519-12
US-08-440-13-13
US-08-440-13-36
US-08-440-13-36
US-08-440-519-10
US-08-440-519-10
US-08-440-519-10
US-08-440-519-10
US-08-440-519-36
US-08-440-519-36
US-08-440-519-36
US-08-440-519-36
                                                                                                                                                                                                                            US-09-046-604-36
US-08-440-549-10
                                                                                                                                                                                                                                                          US-08-850-328-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION:
APPLICATION: 435
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEPHONE: 404-818-3799
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
TENGTH: 28 amino acids
TENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 154; 100.0%; Pred. No. 4
```

Gaps

```
95.5%; Score 147; DB 2; Length 35
92.9%; Pred. No. 5.8e-14;
Live 2; Mismatches 0; Indels
                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungauh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
CITY: Palo Alto
STATE: CA
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPOKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Fabian, Gary R.
REGISTATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US/08/485,500
                                                                                                                                                                                                                                                                 FILING DATE:
CLASSIFICATION 435
PIGNA APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 26, Application US/08485500 Patent No. 5843639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 26
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.97
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 35 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                  ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
CLASSIFICATION: 435
        STREET: 350 TIY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-407-410B-26
                                       CITY: Pa.
STATE: C.
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
US-08-485-500-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 95.5%; Score 147; DB 1; Length 35; Best Local Similarity 92.9%; Pred. No. 5.8e-14; Matches 26; Conservative 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 05-APR-1991
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 26, Application US/08407410B
Patent No. 5843636
GENERAL INFORMATION:
GENERALICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Cristian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 55
CONTRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
                                                                                                                                                                                                                                                                                                  STREET: 350 Cambridge Ave., Suite 250 CITY: Palo Alto COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPOKRNORNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                           1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                   1 PKPOKRNORNINRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DAYS:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                               Sequence 26, Application US/07681703B Patent No. 5443965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE: (415) 324-0880
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 94306
COMPUTER REDABLE FORM:
MEDIUM TYPE: Flore:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     35 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein -07-681-703B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
US-08-407-410B-26
                                                                                                                        RESULT 2
US-07-681-703B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                   ò
                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                Length 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 75;
                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: U.S.A.

21P: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATCHIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/07813
FILING DATE: 19920916
                                                                                                                                                                                                                                                                                                                                                                                                                      CHAD377/A26D
                                              Score 147; DB 5;
Pred. No. 5.8e-14;
2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 147; DB 5;
Pred. No. 1.3e-13;
2; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
                                                                                                                                                     5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: Deblinger & Associates
STREET: 350 Cambridge Ave., Suite 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPOKRNORNTNRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                           APPLICANT: LESNIEWSKI, RICHARD R. APPLICANT: LEUNG, TAT K. TITLE OF INVENTION: HEPATITIS C ASSAY NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Sequence 20, Application US/07681703B
; Patent No. 5443965
                                                                                                                                                                                                                                                                   Sequence 1, Application PC/TUS9207813 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKIP, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4767.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   95.5%;
                                              Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match
Best Local Similarity 92.9°
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             708-937-9556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; TOPOLOGY: linear; MOLECULE TYPE: peptide PCT-US92-07813-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ILLINOIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-07-681-703B-20
        PCT-US91-02370-26
                                                                                                                                                                                                                                   RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 7
                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                            Score 147; DB 2; Length 35;
Pred. No. 5.8e-14;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER FEADABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02370
FLING DATE: 19910405
PRIOR APPLICATION 1435.5
PROOR APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION NUMBER: US 594,854
FILING DATE: 06-APR-1990
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Reyes, Gregory
APPLICANT: Kin, Jungsuh P.
APPLICANT: Kin, Jungsuh P.
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: P.O. BOX 60850
CITY: Palo Alto
FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06 APR-1990
PRIOR APPLICATION DATA: US 594,854
APPLICATION NUMBER: US 594,854
FILING DATE: 09-0CT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REGISTRATION NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
SEQUENCE CHARACTERISTICS:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPOKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPOKKNKRNTNRRPODVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.41
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 26
                                                                                                                                                                                                                                                                                          35 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear
HOLECULE TYPE: protein
US-08-485-500-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 94306
                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

Palo Alto

```
95.5%; Score 147; DB 2; Length 90; 92.9%; Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                         0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 20. Application US/08485500
Patent No. 5843639
CENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Mim, Jungsuh P.
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
CORRESPONDENCE ADDRESS: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
CITY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,500
APPLICALLY

APPLICALLY

ATTORNEY/AGENT INFORMATION:
NAME: Fablan, Gary

REGISTRATION NUMBER: 33,875

REGISTRATION NUMBER: 33,875

REGISTRATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
INFORMATION FOR SEQ ID NO: 20:
LENGTH: 90 amino acids
TOPOLOGY: linear

MOLECULE TYPE: manno acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-407-4108-20

95.5%; Score 147;
                                                                                                                                                                                                                                                                                                                                                                     2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
APPLICATION DATA:
APPLICATION NUMBER: US 594,854
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 20
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
US-08-485-500-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 147; DB 1; Length 90;
Pred. No. 1.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: Patentin Release #1.0, Version #1.25
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/O7/681,703B
CLASSIFICATION 1435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATE: 06-APR-1990
RILING DATE: 06-APR-1990
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INPORMATION:
REFERENCE/DOCKET UNMBER: 33,875
REFERENCE/COCKET UNMBER: 33,875
REFERENCE/COCKET UNMBER: 33,875
TELECOMMULICATION INFORMATION:
TELECOMMULICATION INFORMATION:
TELECOMMULICATION INFORMATION:
TELECOMMULICATION INFORMATION:
TANDELIEPHONE: (415) 324-0880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Repair of APPLICANT: Rim, Jungsuh P.
APPLICANT: Kim, Jungsuh P.
APPLICANT: Meeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes NUMBER OF SEQUENCES: 26
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPOKRNORNINRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 20, Application US/08407410B Patent No. 5843636 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 94306
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 90 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 92.9
Matches 26; Conservative
                             ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    , MOLECULE TYPE: protein US-07-681-703B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
TOPOLOGY: linear
              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 8
US-08-407-410B-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

g

```
Score 147; DB 3; Length 105;
Pred. No. 1.8e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 18, Application US/07681703B

Patent No. 5443965

GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Deblinger & Associates
                                                                                                                                                                                                                                                SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/444,818 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E: Dehlinger & Associates
350 Cambridge Ave., Suite 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPOKRNORNINRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/07/681,703B
FILING DATE: 05-APR-1991
                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14 *MAR-1995
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                     PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US 505,611 FILING DATE: 06-APR-1990
                                       E: Chiron Corporation 4560 Horton Street
                                                                                                                       COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 105 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Harbin, Alisa A. REGISTRATION NUMBER: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
                 CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Co
NUMBER OF SEOUENCES:
                                                       STREET: 4500...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET: 350 Camb
CITY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-07-681-703B-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-444-818-107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                             ö
                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95.5%; Score 147; DB 5; Length 90; 92.9%; Pred. No. 1.5e-13; tive 2; Mismatches 0; Indels
                 95.5%; Score 147; DB 2; Length 90; llarity 92.9%; Pred. No. 1.5e-13; Conservative 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:

SOFTWARE PATENTIN RELEASE #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02370
FILING APPLICATION: 435.5
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: FABILAN, GATY R.

REFERENCE/POCKET NUMBER: 4600-076.41
FELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 107, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT:
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
                                                                                                                                                                                                                                                                                             APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: MocKil, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Mismatches
                                                                                                       1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                              1 PKPOKRNORNINRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PCT-US91-02370-20; Sequence 20, Application PC/TUS9102370; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: Peter J. Dehlinger
STREET: P.O. BOX 60850
CITY: Palo Alto
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 20:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: HER NUMBER OF SEQUENCES: 26 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AMINO ACID
                                          Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PCT-US91-02370-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-444-818-107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                      Query Match
```

ô g

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 147; DB 2; Length 119;
Pred. No. 2.1e-13;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/485,500
             Score 147; DB 2; L
Pred. No. 2.1e-13;
2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Kim, Jungsuh P.
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epit
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. 2.16
2; Mismatches
                                                                                                                                1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                      5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-0CT-1990
ATTORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 33,875
REGISTRATION NUMBER: 33,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 18, Application PC/TUS9102370 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                Sequence 18, Application US/08485500
Patent No. 5843639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: 350 Cambo 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94306
                                                                                                                                                                                                                                                                                                                        US-08-485-500-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US91-02370-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                    á
                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 147; DB 1; Length 119;
Pred. No. 2.1e-13;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM C Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Unugsuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
                       FILING DATE: 09-07-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fablan, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US/08/407,410B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY AGENT INFORMATION:
RESISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 323-8302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 18, Application US/08407410B
Patent No. 5843636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 92.5
These 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: protein US-07-681-703B-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
OGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CITY: Palo
STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-407-410B-18
```

ò

APPLICANT: Reyes, Gregory APPLICANT: Kim, Jungsuh P

```
us-09-491-146a-29.rai
```

```
APPLICANT: Mocckil, Randolph
APPLICANT: Simonsen, Christian C.
YITLE OF INVENTION: Heptilis C Virus Epitopes
NUMBER OF SUGURCES: 26
NUMBER OF SUGURCES: 26
STATE: P. O. BOX 60850
CTTY: Palo Alto
CTTY: Palo Alto
CTTY: Palo Alto
CTTY: Palo Alto
COMPUTE: Palo Alto
```

sarch completed: August 7, 2003, 11:23:52
Job time : 11.6364 secs

ò

```
August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'Cgn2_6/ptcdata/1/pubpaa/US07_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US07_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US06_BWB_PUB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US06_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US07_NEW_PUB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US07_NEW_PUB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US08_PUBCOKB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US10A_PUBCOKB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
'I: 'Cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   451899
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  otal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                       154
1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             451899 segs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                        US-09-491-146A-29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                 Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description	Sequence 2, Appli	Sequence 1, Appli	Sequence 1, Appli	Sequence 2, Appli	Sequence 1. Appli	Sequence 10. Appl	Sequence 1, Appli	Sequence 8 Appli	Sequence 77, Appl	Sequence 14, Appl	Sequence 78, Appl	Sequence 46, Appl	Sequence 60, Appl	Sequence 152, App	152,	
ID	US-10-268-562-2	US-10-098-857B-1	US-10-268-562-1	US-09-916-359-2	US-10-367-677-1	US-09-851-138-10	US-09-758-308-1	US-09-756-875-8	US-09-921-397-77	US-09-851-138-14	US-09-921-397-78	US-09-851-138-46	US-09-851-138-60	US-09-899-046-152	US-09-878-281-152	
DB	15	15	15	6	15	10	6	10	10	10	10	10	10	11	11	
Query Match Length DB	43	122	190	3011	44	74	91	97	103	108	113	137	138	166	166	
Query Match	95.5	95.5	95.5	95.5	9.68	9.68	9.68	89.6	9.68	9.68	9.68	89.6	9.68	89.6	9.68	
Score	147	147	147	147	138	138	138	138	138	138	138	138	138	138	138	
Result No.	~	7	m	4	S	9	7	æ	6	10	11	12	13	14	15	

Sequence 1, Application US/10098857B
Publication No. US20030032031A1
GENERAL INFORMATION:
APPLICANT: GOEDER, JOHN A.
TITLE OF INVENTION: USE OF SPECIFIC ANTIBODY TITERS TO PREDICT HEPATIC

US-10-098-857B-1

RESULT 2

g à

Sequence 42, Appl Sequence 42, Appl Sequence 42, Appl Sequence 44, Appl Sequence 2, Appl Sequence 2, Appl Sequence 12, Appl Sequence 12, Appl Sequence 52, Appl Sequence 54, Appl Sequence 51, Appl Sequence 51, Appl Sequence 50, Appl	00 00 00 00 00 00 00 00 00 00 00 00 00	Length 43; Indels 0; Gaps 0;
11 US-09-899-046-42 11 US-09-899-046-44 11 US-09-878-281-42 10 US-09-878-281-42 10 US-09-878-281-42 11 US-09-104-966-3 11 US-09-104-966-3 10 US-09-851-138-13 10 US-09-851-138-13 11 US-09-899-046-54 11 US-09-899-046-54 11 US-09-899-046-54 11 US-09-899-046-54 11 US-09-899-046-54 11 US-09-899-046-54 11 US-09-899-046-54 11 US-09-899-046-54	-09-878-281-144 -09-973-025-50 -09-995-808-50 -09-995-808-50 -09-995-808-50 -09-995-255-4 -09-95-575-40 -09-95-575-40 -09-95-575-40 -09-95-575-40 -09-95-95-11 -09-1419-20 -09-95-955-1 -09-91-894-3 -09-95-955-1 -09-91-894-3 -09-981-894-3 -09-981-894-3 -09-981-894-3 -09-981-894-3 -09-981-894-3 -09-981-894-3	rus .S%; Score 147; DB 15; .9%; Pred. No. 2.2e-13; e 2; Mismatches 0;
16 138 89.6 169 17 138 89.6 169 20 138 89.6 169 21 138 89.6 182 22 138 89.6 182 24 138 89.6 182 25 138 89.6 319 26 138 89.6 319 27 138 89.6 319 28 138 89.6 319 29 138 89.6 319 20 138 89.6 319 21 13 89.6 319 22 13 89.6 319 23 13 89.6 319	18 89.6 31 18 89.6 86 18 89.6 86 18 89.6 86 18 89.6 289 18 89.6 289 18 89.6 301 18 89.6 301 10 000000000000000000000000000000000	; TYPE: PRT ; ORGANISM: Hepatitis C virus US-10-268-562-2 Query Match Best Local Similarity 92.99 Matches 26; Conservative

```
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AND HAVE INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/10/367,677
CURRENT FILING DATE: 2003-02-19
PRIOR FILING DATE: 1999-09-07
PRIOR PELICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442
PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-05
NUMBER OF SEQ ID NOS: 11
SSOFTWARE: PATENTIN VOT: 2.1
                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                  Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.6%; Score 138; DB 15;
85.7%; Pred. No. 3.8e-12;
                                                                                                                  Score 147; DB 9;
Pred. No. 2e-11;
2; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                   1 PKPQKRNORNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                         1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1, Application US/10367677 Publication No. US20030118604A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGENTS
                                                                                                                  95.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
ORGANISM: Hepatitis C virus
                                                                                                               Query Match 95.5
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 85.7
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: JOLIVET, MICHEL
                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
        LENGTH: 3011
TYPE: PRT
ORGANISM: Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
US-09-851-138-10
                                                                        US-09-916-359-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-367-677-1
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1, Application US/10268562

Folication No. US20030108563A1

GENERAL INFORMATION:
GENERAL SEPERATION:
TITLE OF INVENTION:
APPLICATION WINBER: US/10/268,562
CURRENT FILING DATE: 2002-10-10
PRIOR PELLING DATE: 2001-11-07
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
        FAILURE IN PEOPLE INFECTED WITH HEPATITIS C VIRUS
                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                 Length 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Veronique Barban
TITLE OF INVENTION: VERCINE COMPOSITION FOR PREVENTING OR
TITLE OF INVENTION: TREATING C HEPATITIS
FILE REFERENCE: PMCF97-03A
CURRENT APPLICATION NUMBER: US/09/916,359
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: 09/388,874
PRIOR FILING DATE: 1999-09-02
PRIOR FILING DATE: 1997-03-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                         0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 147; DB 15;
Pred. No. 1.1e-12;
? Mismatches 0;
                                                                                                                                                                                                                                                                                                                                             Score 147; DB 15;
Pred. No. 6.6e-13;
FILE REFERENCE: 1300-1800.01
CURRENT APPLICATION WUMBER: US/10/098,857B
CURRENT APLICATION WUMBER: US/10/098,857B
CURRENT FILING DATE: 2002-10-15
PRIOR APPLICATION WUMBER: 09/616,823
PRIOR FILING DATE: 2000-07-14
PRIOR FILING DATE: 1999-07-15
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PATENTIN VET. 2.1
                                                                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPOKRNORNINRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPOKRNORNINRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      75-09-916-359-2
; Sequence 2, Application US/09916359
; Patent No. US20020034734A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                           Query Match 95.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                         ; ORGANISM: Hepatitis C virus US-10-098-8578-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; TYPE: PRT; ORGANISM: Hepatitis C virus US-10-268-562-1
                                                                                                                                                                                                                                        LENGTH: 122
                                                                                                                                                                                                                  SEQ ID NO 1
                                                                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 77, Application US/09921397

Patent No. US20020151484A1

GENERAL INFORMATION:

APPLICANT: HYBRIGENICS

TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof

FILE REFERENCE: A8409A - JAZ

CURRENT PAPFLICATION NUMBER: US/09/921,397

CURRENT FILING DATE: 2001-08-03

PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 97;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: PatentIn Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 138; DB 10;
Pred. No. 8.9e-12;
             Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APIRE, TATLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 7701-E Columbia Square
STREET: 555 13th Street, N. W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89.6%; Score 138;
85.7%; Pred. No. 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 85.7<sup>3</sup>
Matches 24; Conservative
                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         , MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                           CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                     u.s.
                                                                                                                                                                                                                                 STATE: D. C.
                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE
                                                                                                                                                                                                                                                                       ZIP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                     COUNTRY:
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT FILING DATE: 2001-01-10
CURRENT FILING DATE: 2001-01-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 74;
                                                                                                                                       COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.6%; Score 138; DB 9; L6
85.7%; Pred. No. 8.3e-12;
Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match 89.6%; Score 138; DB 10; Best Local Similarity 85.7%; Pred. No. 6.7e-12; Matches 24; Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                    PATURA APPLICATION NUMBER: 08/836,075
FILING DATE: «UNKNOWN)
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jul 1995
ATTORNEY AGENT INFORMATION:
REGISTRATION NUMBER: 29,775
REGISTRATION NUMBER: 29,775
INFORMATION FOR SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ) MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                             COUNTRY: USA
ZIP: 77210-443
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1, Application US/09758308
Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PatentIn version 3.0
  STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 85.77
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION NUMBER:
PRIOR FILING DATE: 1999-
NUMBER OF SEQ ID NOS: 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBERS
SOFTWARE: PR
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

RESULT 8

```
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROFHYLACTIC, THERAPEUTIC AND DIAGNO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                  Length 113;
                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89.6%; Score 138; DB 10;
85.7%; Pred. No. 1.3e-11;
iive 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 6.0 / ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                              Score 138; DB 10
Pred. No. 1e-11;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
        FILE REFERENCE: B4809A - JAZ
CURREWT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: EP 00402225.7
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPOKRNORNINRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 46, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: USA
21P: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                              89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                               ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 85.77
Matches 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 85.7 tes 24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-851-138-46
                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                          STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from
TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
TITLE OF INVENTION: applications thereof
                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
        Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC_compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
  Score 138; DB 10;
Pred. No. 9.5e-12;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 89.6%; Score 138; DB 10; Best Local Similarity 85.7%; Pred. No. 9.9e-12; Matches 24; Conservative 3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                           1 PKPOKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                            Sequence 14, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: USA
21P: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                          AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: P.O. BOX 4433
ch 89.6%;
1 Similarity 85.7%;
24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CITY: HOUSTON
STATE: TEXAS
                         Best Local Similarity
                                                                                                                                                                                      RESULT 10
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-921-397-78
  Query Match
                                          Matches
                                                                                     à
                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
```

Gaps

g

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                        Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 166;
                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                   Score 138; DB 11;
Pred. No. 1.6e-11;
3; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 11;
Pred. No. 1.6e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                     1 PKPOKRNORNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7, 2003, 12:01:13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 152, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89.6%;
                                                                                                                                                                                                                                                                                     89.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24; Conservative
                                                                                                                                                                                                                                                                                                                            24; Conservative
                                                                                                                                                                                              TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; MOLECULE TYPE: protein US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: August
Job time : 14.3636 secs
                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-878-281-152
                                                                                                                                                                                                                                                                                     Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                          TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      of hepatitis C virus diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICTOSOFT WORD 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89.6%; Score 138; DB 10; Length 138;
llarity 85.7%; Pred. No. 1.3e-11;
Conservative 3; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
OPERATING SYSTEM: PC-DOS/MS-DOS
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: 08/836,075
FILING DATE: «Unknown»
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQKRNQRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: New sequences TITLE OF INVENTION: genotypes for NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 152, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                   Sequence 60, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                   APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                        STUYVER, LIEVEN
                                                                                                                                                                                                                                                                    AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 24; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-899-046-152
                                                                                 RESULT 13
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1-09-851-138-60
```

ö

Gaps

g

```
AAR30689
AAR54065
AAW37380
AAW66083
AAY26952
AAY26952
AAY26953
AAX29535
 1220013
\alpha
110
110
110
110
110
110
                                          August 7, 2003, 11:05:37; Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
                                                                                                                                                                                       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                         tal number of hits satisfying chosen parameters:
                                                                          154
1 PKPORKTKRNAHRRPODVKFPGGGOIVG 28
                                                                                                              1107863 segs, 158726573 residues
                               OM protein - protein search, using sw model
                                                                                                                                                                      Listing first 45 summaries
                                                                                                  Gapop 10.0 ', Gapext 0.5
                                                                                                                                                                                    A_Geneseq_19Jun03:*
                                                                                                                                                                 Maximum Match 100%
                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                          Post-processing: Minimum Match 0%
                                                                   US-09-491-146A-28
                                                                                            BLOSUM62
                                                                                                                                                                                                                                                                              Scoring table:
                                                                         Perfect score:
Sequence:
```

Database :

Searched:

Run on:

Title:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	NC mosaic protein	NC mosaic protein	Hepatitis C virus	NC_mosaic protein	Blood transmiscibl	HCV-S1 full-length	CN14 fragment of H	HCV capsid peptide	HCV capsid peptide
D		AAY06683	AAR58593	AAY06675	AAR53417	AAE20477	AAW06487	AAR30687	AAR30688
BB	20	20	15	20	15	23	16	14	14
% Query ore Match Length DB I	28	28	312	28	3010	3010	36	38	38
& Ouery Match	100.0	96.1	95.5	93.5	93.5	93.5	92.9	92.9	92.9
Score	154	148	147	144	144	144	143	143	143
Result No.	-	7	m	4	ις	Q	7	80	6

HCV capsid peptide Non-A, non-B hepat Hepatitis C virus Hepatitis C virus Hepatitis C virus Hepatitis C virus Human hepatitis C HUMAN hepatitis C HUMAN CORE-envelope HCV CORE-envelope HCV CORE-envelope HCV CORE-envelope HCV CORE-envelope HCV CORE-envelope HCV CORE-envelope Non-A, Non-B Hepat PROTOTIE BASE ANTI-HCV ANTIBODY PROTOTIE DASE ANTI-HCV ANTIBODY PROTOTIE DASE ANTIGEN PHCAIOI. HCV fragment 1 / I HCV fragment 2 / I BRANCHED PHCAIOI. HCV Antigenic epitope HCV fragment C virus FCOTEIN ENCOMED BY HEPATITIS C virus HCV bait polypepti HCV bait polypepti HCV capsid CORE pr HCV CORE-ENVELOPE	TENTS.  TOTE COMPANY TOTE TO THE STATE TO TH
AAR30689 AAR34065 AAW37380 AAW66083 AAY26952 AAY294410 AAY29535 AAR29536 AAR29536 AAR29536 AAR29536 AAR29542 AAR2954121 AAR40978 AAR66619 AAR40978 AAR66619 AAR40978 AAR6663 AAR40978 AAR6663 AAR1655 AAR40978 AAR6663 AAR40978 AAR6663 AAR1655 AAR40978 AAR6663 AAR40978 AAR6663 AAR6663 AAR69530 AAR6663	ment b; nucliligate l'ilgate plural cof a an inc
1005 220 0 10 10 10 10 10 10 10 10 10 10 10 10	rentry)  t entry)  ingenic p  ligase as  lease as  -US17385  -US17
\$	standard;  99 (firs protein otein; an on endonuv C virus. Al. 99. 98; 98% 97; 97U s DEPT HE c Khudya c Codefil/i
	1 700678; AY06678; AY06678; C mosaic DSAic pr PSAICTICTI PPAILITIS 1-AG-19
	AAAYO665 AAAYO665 AAAYO665 AAAAYO665 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

the sensitivity,

셤

```
artificial gene and the resulting mosaic protein improve the sensitive spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAYO6673-683 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP06225770-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUL-1992;
                                                                                                                                                                                                                                                                                                                                                                                                               09-MAY-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-AUG-1994
                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                            AAR58593;
                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOFU )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SANW )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY06675
ID AAY0
                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                                                         AAR58593
                                                                                                                                                                                                                                                                                                                                                                             88888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                            ö
                                            The invention relates to a mosaic protein, comprising a plurality of annologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention relates to a mosaic protein, comprising a plurality of monologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted provided. The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for
                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                       Length 28;
                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                         20;
                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 154; DB 20; 100.0%; Pred. No. 2.3e-14; tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           peptides from different genotypes of a species detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                               NC mosaic protein amino acid fragment K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY06683 standard; Protein; 28
              Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
les 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                     28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fields HA,
                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY06683;
                                                                                                                                                                                                                                                                                                                                                         ery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim
                                                                                                                                                                                                                                                                                                                                                                          ast Loc
Matches
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAQ70543 is a fragment of hepatits C virus (HCV) or non-B hepatits virus (NAMBH) core and envelope one structural regions, encompassing base pairs 349-1285. It codes for AARS8593 an antigen to a structural region of the HCV virus, which can be used in the diagnosis of NANBH patients and the detection of HCV carriers.
                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                               Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for diagnosis of NANBH and detection of HCV \,
                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                      Score 148; DB 20;
Pred. No. 1.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 147; DB 15;
Pred. No. 2.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO
                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus 349-1285 fragment antigen
                                                          Pred. No. 1.66
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 2.66
1; Mismatches
                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SANWA KAGAKU KENKYUSHO CO. TONEN CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 1; Page 18; 22pp; Japanese.
                                                                                                                                                                                                                                                                  AAR58593 standard; Protein; 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY06675 standard; Protein; 28
                                      96.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 95.5%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   93JP-0193104.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (ROKU-) KOKUSAI SHIYAKU KK
                                                          Best_Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1994-298800/37.
N-PSDB; AAQ70543.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             312 AA;
28 AA;
```

17-JUN-1999

AAY06675;

04-MAR-1999

```
Leu
                                                                                                                                                                                                                                           G1y
Location/Qualifiers
                                                                                                                                                                                                                                       label Ser, Arg,
                                                                                                                                                                                                                                                                                                           Pro,
                                                                                                                                                                                                                                                                                                                                               Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             _{\rm Ile}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Met
                                                                                                                                                                                                       label- Gly, Asp
                                                                                                                                                                                                                                                                                                                                                                               Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          G1u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Thr, Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label= Cys, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      abel- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label - Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label= Lys,
                            label- His,
                                                                                                                                   label- Asp,
                                                                                                                                                                                                                                                                                                           label- Gln,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 label- Leu,
                                                                                                label- Leu,
                                                                                                                                                                                                                                                                         label= Thr,
                                                                                                                                                                                                                                                                                                                                             label= Lys,
                                                                                                                                                                                                                                                                                                                                                                               label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                 label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label = Met,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label= Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label= Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label- Val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label - Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label= Ser,
                                                                                                                                                                     label- Phe,
                                                                label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 1455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 2917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2989
                                                                                                                                                                                                                                                                                         Misc-difference 405
                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 759
                                                Misc-difference 226
                                                                                                                                                                                     Misc-difference 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 1
 Key
Misc-difference
                                                                              Misc-difference
                                                                                                                 Misc-difference
                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosalc protein and the artificial mosalc protein are useful for detecting a hepatitis infection in an individual. The mosalc gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosaic protein improve the sensitivity, and antigen specificity of enzyme immunosacitylity, and antigen specificity of enzyme sequences AXY06673-683 represent amino acid sequence of each monomer comprising the NC mosaic protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR; C100 antibody; HCV RNA; NS5 region.
                                                                                               Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The invention relates to a mosaic protein, comprising a plurality of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                              New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species - useful for detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93.5%; Score 144; DB 20; Length 28; 92.9%; Pred. No. 5.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NC mosaic protein amino acid fragment C.
                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Non-A, non-B hepatitis virus
                                                                                                                                                                                                                                                     98WO-US17385.
                                                                                                                                                                                                                                                                                      97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              est Local Similarity 92.9
Atches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                        Fields HA, Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-204671/17.
                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28 AA;
                                                                                                                                                                                 WO9910506-A1
                                                                                                                                                                                                                                                    21-AUG-1998;
                                                                                                                                                                                                                                                                                      25-AUG-1997;
```

17-JAN-1995

AAR53417;

RESULT 5 AAR53417

Sequence

**Nery Match** Matches

ö g

```
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAT45055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             core region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                             cassette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW06487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
  ####X & X O O O O O O O O O O O O O O X &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME OF A PART O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                                                                            This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQ63500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were C100 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed from 23 clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                          Blood-transmissible non-A non-B hepatitis virus DNA - used for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93.5%; Score 144; DB 15; Length 3010; 92.9%; Pred. No. 6.9e-11; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKPORKTKRNTYRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MOLE-) INST MOLECULAR & CELL BIOLOGY (EHRL/) EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                   Claim 1; Page 8-20; 22pp; Japanese.
                   /label- Tyr, Cys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                               detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24-JUL-2000; 2000US-220248P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-JUL-2001; 2001WO-IL00669
                                                                                                                                         92JP-0051885
                                                                                                                                                                                  92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ran YH, Lim SP, Lim SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 92.9
nes 26; Conservative
  Misc-difference 2990
                                                                                                                                                                                                                                                              WPI; 1994-163130/20.
N-PSDB; AAQ63499.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-280605/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                          (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAD33038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200208447-A2.
                                                           JP06105690-A
                                                                                                   19-APR-1994.
                                                                                                                                         10-MAR-1992;
                                                                                                                                                                                  10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31-JAN-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAE20477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAE20477
ID AAE
XX
AC AAE
XX
DT 01-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
The invention relates to nucleic acid construct which comprises an expression cassette including a first polynucleotide region including a 5 non-coding region (NCR) sequence of an RNA virus and at least an N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3 untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus and a thirst polynucleotide region encoding a reporter molecule, flanked by first and second polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug
Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of HCV infection and to raise antibodies against it.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                resistance of an RNA virus. The present sequence is Hepatitis \check{\mathbf{C}} virus (HCV) isolate HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Detection of hepatitis C virus - using oligopeptide fragment of HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    93.5%; Score 144; DB 23; Length 3010; 92.9%; Pred. No. 6.9e-11; 1. Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                        Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW06487 standard; Protein; 36 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CN14 fragment of HCV core region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93JP-0156026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93JP-0156026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (IMMO ) IMMUNO JAPAN KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1995-047903/07.
```

AAR30687;

UXXXXEEX

RESULT 8

ò g

```
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific antibodies.
                                                                                                                                         Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ch 92.9%; Score 143; DB 14; Length 38; 1 Similarity 92.9%; Pred. No. 1.1e-12; 26; Conservative 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Updated on 25-MAR-2003 to correct PN fleld.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR30689 standard; peptide; 38 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Todd JA;
                                                                                                                                                                                                                                                                                                                                                                                                                        (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                  91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                92WO-US03635
                                                                      (first entry)
                                                                                                        HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV capsid peptide No. 25.
                                                       (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38 AA;
                                                                                                                                                                                                                                                       W09222571-A1.
                                                                                                                                                                                                                                                                                                                                                                  13-JUN-1991;
20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                29-APR-1992;
                                                 25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                            23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jolley ME,
               AAR30688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR30689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
AAR30689
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The sequences given in AAR30665-89 represent fragments of the
             Length 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 38;
                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                 1;
           Score 143; DB 16;
Pred. No. 1e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.9%; Score 143; DB 14; 92.9%; Pred. No. 1.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRNAHRRPODVKFPGGGGIVG 28
                                                                                       1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                          Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR30688 standard; peptide; 38 AA.
                                                                                                                                                                                                                   Ą
                                                                                                                                                                                                                 AAR30687 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BAXT ) BAXTER DIAGNOSTICS INC
             92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92WO-US03635
                                                                                                                                                                                                                                                                                          (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                               HCV capsid peptide No. 23
                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1993-018073/02
     Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            virus – for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W09222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20-JUN-1991;
                                                                                                                                                                                                                                                                                        25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jolley ME,
```

ö

Gaps

; 0

Sequence

;888888888**%**8

RESULT 9
AAR30688
ID AAR3

ð g 9

ö

```
The present sequence represents a Hepatitis C virus (HCV) protein sequence from the disclosure of the present specification. The present specification describes a chimeric HCV peptide antigen which comprises at least 2 peptide epitope regions from the HCV polypeptide core region, 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS3 region and at least 2 peptide epitope regions from the NS4 region. The antigen binds specifically with an antibody produced by a human infected by HCV. The peptide can detect a wide range of HCV infections with high sensitivity.
                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  peptide antigen derived from hepatitis C virus protein detecting HCV infections
                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus; HCV; chimeric; antigen; detection; core region; epitope; NS3; NS4; infection.
                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
               DNA coding a Non-A, non-B hepatitis virus antigen - useful for
detecting HCV within serum
                                                                                Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NSI, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.9%; Score 143; DB 19;
llarity 92.9%; Pred. No. 1.2e-12;
Conservative 1; Mismatches 1;
                                                                                                                                                                  Score 143; DB 15;
Pred. No. 1.1e-12;
1; Mismatches 1;
                                                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Page 24; 30pp; Japanese.
                                                      Disclosure; Page 11; 22pp; Japanese
                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus C-1 protein 1-43.
                                                                                                                                                                                                                                                                                                                     AAW37380 standard; Protein; 43 AA.
                                                                                                                                                                Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97JP-0027015.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96JP-0024045
                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1998-022248/03.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
les 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JP09278794-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New chimeric - useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .0-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                           11-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-OCT-1997.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                              AAW37380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                        AAW37380
  ð
                                                                                                                                                                                                                                                DD
                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                           The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capaid protein of the virus and themselves contain epitope groups. These peptides can be used in immunoassays for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and antibodies.
                                                                                                                                                                                                                   Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Non-A, non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; core; ENV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 143; DB 14; Length 38;
Pred. No. 1.1e-12;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Non-A, non-B hepatitis virus coreI region fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (SANW ) SANWA KAGAKU KENKYUSHO CO.
(TOFU ) TONEN CORP.
(TOKR-) 2H TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                                                                           (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR54065 standard; Protein; 38
                                                                                                                                  (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.9%;
92.9%;
                                                                                         91US-0714471.
91US-0718052.
                                                                92WO-US03635.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                              Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1994-205028/25.
N-PSDB; AAQ64067.
                                                                                                                                                                                         WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                       38 AA;
                                                                                          13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JP06141870-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-MAR-1992;
                                                                29-APR-1992;
                                                                                                          20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14-FEB-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      12-MAR-1992;
                                      23-DEC-1992.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-MAY-1994
                                                                                                                                                            Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR54065;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR5406
88888888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
N-terminus of the hepatitis C virus core protein. The peptides are used to generate monoclonal antibodies or antibody fragments specific for hepatitis C virus (HCV) core protein. The monoclonal antibodies are used for early diagnosis of HCV infections, especially by two-antibody
                                                                                                                                                                                                                                                                                                                                                                      useful for detecting and/or quantifying hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polypeptides that bind to anti-hepatitis C virus antibodies, useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human; hepatitis C virus; HCV; immunogen; antigen; antibody; virucide;
hepatotropic; anti-inflammatory; virus detection; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Peptides AAY26949-Y26955 represent peptide epitopes derived from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human hepatitis C virus core protein N-terminus, residues 2-45.
                 Epitope; hepatitis C virus; core protein; monoclonal antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Jolivet-Reynaud C;
                                                                                                                                                                                                                                                                                              Jolivet M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; DB 20;
1.2e-12;
                                                                                                                                                                                                                                                                                              Paranhos BC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 143; DB Pred. No. 1.2e 1; Mismatches
                                  diagnosis; infection; sandwich immunoassay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ź
                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 3; Page 10; 19pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY94410 standard; peptide; 44
                                                                                                                                                                                                                                                                                              Yvon S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                               98FR-0003087
                                                                                                                                                                                                                     98FR-0003087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99WO-IB01933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0196155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11-SEP-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ď
                                                                                                                                                                                                                                                                                                                                                                  Monoclonal antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dalbon
                                                                                                                                                                                                                                                                                                                                                                                       C virus core protein
                                                                                                                                                                                                                                                                                              Piga N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sandwich immunoassay
                                                                                                                                                                                                                                                                                                                               WPI; 1999-530397/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       API; 2000-411934/35
                                                                                                                                                                                                                                                          (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             44 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Н
                                                                                                                                                                               09-MAR-1998;
                                                                                                                                                                                                                     09-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20-NOV-1998;
                                                                                                       PR2775690-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-JUN-2000.
                                                                                                                                          10-SEP-1999
                                                                                                                                                                                                                                                                                            Jolivet RC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dalbon P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY94410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY94410
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The invention relates to a peptide, which is recognised by antibodies against amino acids 2-45 at the N-terminus of the core (or nucleocapsid) possibly to the period of hepatitis C virus (HCV), or its variants. The peptide has a tertiary structure consisting of two alpha-helical fragments, almost compendicular to each other in space, connected by a junction peptide. Excluded are all proteins and peptides comprising, or consisting of, the N-terminal part of p21 starting from amino acid to x 2. Also new are cc (1) monoclonal or polyclonal antibodies produced using the peptide as an immunogen and (2) complex consisting of the peptide specifically cound to some other molecule, particularly peptide or nucleotide concentrating antibodies and (ii) for detecting and capments or functionalised aromatic compounds. The peptide is used (i) as immunogen for generating antibodies and (ii) for detecting and capments formation). Antibodies are used correspondingly to detect HCV or related antigens. The peptide and antibodies may also be used to treat or prevent HCV infections. The present sequence represents the N-terminus contracts of the core p21 protein of hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New peptide from the N-terminus of hepatitis C virus p21 protein containing the immunodominant epitope - and related antibodies, u for diagnosis, treatment and prevention of hepatitis C infection
                                                                                                                                                                           antigenic; nucleocapsid; p21 protein; alpha-hellcal; immunogen; antibody; hepatitis C virus; epitope; N-terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.9%; Score 143; DB 19; Length 44; 92.9%; Pred. No. 1.2e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Penin F;
                                                                                                                                          Hepatitis C virus p21 protein N•terminal fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus Core protein amino acids 2-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ladaviere L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW66083 standard; peptide; 44 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 16; 37pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lacoux X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY26952 standard; peptide; 44
                                                                                                                                                                                                                                                                                                                                               98WO-FR00442.
                                                                                                                                                                                                                                                                                                                                                                                   97FR-0002878
                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Jolivet M,
                                                                                                                                                                                                                                                                                                                                                                                                                      (INMR ) BIO MERIEUX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1998-495793/42
                                                                                                                                                                                                                                     hepatitis c virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44 AA;
                                                                                                                                                                                                                                                                                                                                                                                   05-MAR-1997;
                                                                                                      16-NOV-1998
                                                                                                                                                                                                                                                                                                                                               05-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-DEC-1999
                                                                                                                                                                                                                                                                                                            11-SEP-1998,
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dalbon P,
                                                                 AAW66083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY26952;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY26952
ID AAYX
XX AAY
AC AAYX
XX AAY
```

SXS

g ò

ö

Gaps

diagnosing and preventing hepatitis C infections -

The present peptide, designated S42G, corresponds to residues 2 to 45 of the N-terminus of the human hepatitis C virus (HCV) core protein. It is an immunodominant region containing conformational type epitopes and linear type epitopes. It manifests an immunoreactivity with all the sera of individuals or blood samples infected with HCV and which possess substituted for homologous amino acros protein. An amino acid may be substituted for homologous amino acids and side chains and peptide bonds may be modified. For example, L-amino acids may be replaced by D-amino sequence and its antiqenic derivatives may be used for detection of hepatitis C virus and for raising antibodies against the virus. Claim 1; Page 42; 50pp; English. F×8×55555555555588

44 AA; Sequence

Gaps ó 92.9%; Score 143; DB 21; Length 44; 92.9%; Pred. No. 1.2e-12; tive 1; Mismatches 1; Indels Query Match 92.99

Pest Local Similarity 92.99

tches 26; Conservative

ö

1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28

4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31

g à

Search completed: August 7, 2003, 11:14:07 Job time : 38.6364 secs

```
GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 11:05:41; Search time 9.54545 Seconds

(without alignments)
282.095 Million cell updates/sec
154
Perfect score: 154
Sequence: 1 PRORKTKRNAHRRPQDVKFPGGGQIVG 28
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 283308 seqs, 96168682 residues
:al number of hits satisfying chosen parameters: 283308
```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR\_76:\* 1: pir1:\* 2: pir2:\* 3: pir3:\* 4: pir4:\* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	депоше ро												_								prc	polyprotein - hepa	ı	genome polyprotein	genome polyprotein			41	genome polyprotein
SUMMARIES																														
SUM	ΩI		541353	841355	S41357	541348	S41371	S41341	S41370	841369	S41368	841342	S41344	841350	841354	841345	841347	S41343	S41346	S21471	S12707	PC1284	J01925	JQ1926	JH0711	S19876	S18031	S18032	PN0677	JQ0883
	DB	-	~	~	~	~	~	~	~	~	~	~	7	~	7	~	~	ď	7	~	~	a	~	~	~	~	~	N	7	7
	Query Match Length		108	108	108	108	112	112	114	114	114	115	115	115	115	115	115	115	118	369	441	513	520	523	550	782	782	782	787	874
di	Query		ς.	ď	ä	ς.	ς.	92.9	ď	~;	ä	ď	'n	ä	ď	ä	ď	ς.	ς.	ď	ä	ä	'n.	ά.	ä	ä	ä	ď	ď	ĸ.
	Score	14	4	4	143	4	4	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143	143
	Result No.	1	7	m	4	Ŋ	9	7	æ	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.

genome polyprotein polypeptide - hepa	genome polyprotein genome polyprotein				genome polyprotein			genome polyprotein				
JQ0881 PC2219	GNWVTC	S18030 GNWVCH	S40770	JC5620	JQ1303	GNWVJ8	S21336	S41356	S41359	S41358	841351	841349
00			-		Н		N	7	7	~	~	0
874 876	3010 3010	3010	3011	3014	3033	3033	88	108	114	114	115	115
92.9	92.9 92.9	92.9	92.9	92.9	92.9	92.9	91.6	90.3	90.3	90.3	90.3	90.3
143	143 143	143	143	143	143	143	141	139	139	139	139	139
30 31	33 33	34	36	37	38	39	40	41	42	43	44	45

## ALIGNMENTS

RESULT 1
genome polyprotein - hepatitis C virus (strain JT) N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nons protein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C.Species: nepailis C.Vilus C.Barte: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001 C.Aarnession: AAF573
R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hiji Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carr A;Reference number: A45573; MUID:92295714; PMID:1318627
A; Accession; A455/3 A; Status: preliminary
A;Cross-references: GB:D11168; GB:D01171; NID:9221612; PIDN:BAA01943.1; PID:922161 A;Experimental source: HCV-JT
A;Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207) C;Superfamily: hepatitis C virus genome polybrotein
C: Keywords: ATP: glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; F:2-115/Product: canaid nantein ( #thatis products of process)
F;116-191/Product: envelope protein M #status predicted <epm></epm>
F;192-389/Product: major envelope protein E *status predicted <mee> F:390-729/Product: nonstructural protein NS1 *status predicted <ns1></ns1></mee>
F;730-1006/Product: nonstructural protein NS2 #status predicted <ns2></ns2>
F.130', 1917/Figurate. inspectfulli "status preulecto (No.) F.130'-123'/Region: nucleotide-binding motif A (P-loop) F.1317-137'/Fordion-identide. District of the control o
F.1316-1319/Region: Dexy motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <n4a> F:1863-2013/Product: nonstructural protein NS4b #status predicted <n4b></n4b></n4a>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <ns5></ns5>
Query Match 93.5%; Score 144; DB 1; Length 3010; Best Local Similarity 92.9%; Pred. No. 1.1e-11; Matches 26; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
DD 5 PKPQKTKRNTYRRPQDVKFPGGGIVG 32
RESULT 2
genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
C:Species: hepatitis C virus A:Variety: genotype 2. N2
C. Accession: S41353

Gaps ö

Indels

```
Similarity 92.9%; Pred. No. 6.2e-13; 26; Conservative 1; Mismatches 1;
                                                                                                                                                            1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPODVKFPGGGOIVG
                          Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
S41341
                                                                                                                                                            ô
                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Saturation of the protein - hepatitis C virus (genotype 2, N6) (fragment)
N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 2, N6
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accasaion: S41357
R.Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Reference number: S41341
A.Accasaion: S41357
A.Accasaion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 84135
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: $41341
A.Accession: $4353
A.Molecule type: genomic RNA
A.Residues: 1-108 <-War
A.Residues: 1-108 <-War
A.Cross-references: EMBL:229456
A.Cross-references: EMBL:229456
A.Cross-references: EMBL:229456
A.Cross-references: EMBL:20456
A.Cross-references: C virus genome polyprotein
C.Superimental Source: genotype 2, N2
A.Cross-references: core protein; core protein; core protein
F:1-108/Product: core protein #status predicted <-MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.9%; Score 143; DB 2; Length 108; 92.9%; Pred. No. 6.2e-13; ive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143; DB 2; Length 10
Pred. No. 6.2e-13;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A)Experimental source: genotype 2, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Reference number: S41341
A; Accession: S41355
A; Molecule type: genomic RNA
A; Residues: 1.108 < VAN>
A; Cross-references: EMBL: 229458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 92.5
les 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
Matches 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

Length 108;

DB 2;

Score 143;

95.98;

Query Match

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                 %.Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Varlety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41371
R;Van Doorn, L.J.; Rieter, G.E.M.; Brouwer, J.T.
A;Danitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
                                                                                                                                                                                                                              ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                           S
                                                                                                                        C;Accession: S41348
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1
A;Reference number: S41341
A;Accession: S41348
A;Mclession: S41348
A;Residues: 1-108 <VAN>
                                                                                                                                                                                                                                                                                                                                                         A. Cross-references: EMBL:229451
A. Experimental source: genotype 1, N6
C. Superfamily: hepatitis C virus genome polyprotein
C. Keywords: capsid protein; core protein; polyprotein
F:1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A) Experimental source: genotype 5, N5 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 92.9%; Score 143; DB 2; Best Local Similarity 92.9%; Pred. No. 6.2e-13; Matches 26; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 143; DB 2;
Pred. No. 6.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Accession: S41371
A;Molecule type: genomic RNA
Residues: 1-112 <7ANN>
A;Cross-references: EMBL:229474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N;Contains: core protein
C;Species: hepatitis C virus
```

Fri Aug

```
C; Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                       Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Reference number:
A; Accession: S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S41368
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ad 370 genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
Noontains: core protein
Noontains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 5, N4
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541370
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Reference number: 541341
A; Accession: 541370
A; Accession: 541370
A; Accession: 541370
A; Residues: 1-114 < VAN>
A; Residues: 1-114 <
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N;Contains: core protein
C;Species: hepatitis C virus
A;Variety; genotype 5, N3
C;Sate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 841369
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Accession: 841369
A;Molecule type: genomic RNA
A;Residues: 1-114 <VAN>
A;Kross-references: EMBL:229472; NID:9443906; PIDN:CAA82610.1; PID:9443907
A;Experimental source: genotype 5, N3
C;Superfamily: hepatitis C virus genome polyprotein
A;Variety: genotype 1, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 4JJ341
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: 54J341
A;Molecule type: genomic RNA
A;Residues: 1-112 <VANA
A;Cross-references: EMBL:229444; NID:g443850; PIDN:CAA82582.1; PID:g443851
A;Experimental source: genotypel, N1
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capaid protein; core protein; polyprotein
F;1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 143; DB 2;
Pred. No. 6.4e-13;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143; DB 2;
Pred. No. 6.5e-13;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HILLIHILL : HILLIHIHIHIHIHIH PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 26; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                     ö
                                                                                                                                                                                                                                                                                                                                               A; Variety: genotype 5, N2
C; Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 54136
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A, Cross-references: EMBL: 229445; NID: 9443852; PIDN: CAA82583.1; PID: 9443853
A; Experimental source: genotype 1, N10
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: Z29471; NID: g443904; PIDN: CAA82609.1; PID: g443905
A; Experimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-114/Product: core protein #status predicted <KAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                               polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C.Accession: S41342
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by A;Reference number: S41341
A;Accession: S41342
  Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143; DB 2;
Pred. No. 6.5e-13;
1; Mismatches 1.
                         Pred. No. 6.5e-13;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 143; DB 2;
Pred. No. 6.6e-13;
     Score 143; DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 1, N;Contains: core protein S,Species: hepatitis C virus A;Vaziety; genotype 1, N10 C;Date: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                   28
                                                                                                                             28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                   1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.9%;
92.9%;
1larity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26; Conservative
                                                                                                                                                                                                                                                                                                 N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
```

```
Search completed: August 7, 2003, 11:21:48
Job time : 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
S41345
                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
                N; Contains: core protein
C; Species: hepatitis C virus
A; Variety; genotype 1, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41344
R; Van Doorn, L, J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Reference number: S41341
A; Residues: 1-115 <-VAN>
A; Residues: 1-115 <-VAN>
A; Residues: 1-115 <-VAN>
A; Cross references: EMBL: 220447; NID: 9443856; PIDN: CAA882585.1; PID: 9443857
A; Experimental source: genotype 1, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <-MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (genotype 1, NB) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, NB
C;Date: 19-May-1594 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41350
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N;Contains: core protein
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N3
C;Date: 19-May_1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A)Cross-references: EMBL:229453; NID:9443868; PIDN:CAA82591.1; PID:9443869
A)Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Accession: S41334
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Reference number: S41341
A; Reference number: S41354
A; Accession: S41354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö

    hepatitis C virus (genotype 1, N2) (fragment)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.9%; Score 143; DB 2; Length 115, 92.9%; Pred. No. 6.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143; DB 2;
Pred. No. 6.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 92.9 ses 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: S41350
A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Juery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
S41354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N3
C; Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Reference number: $41341
A; Accession: $41345
A; Accession: $41345
A; Molecule type: genomic RNA
A; Residues: 1-115 < VAN>
A; Cross-references: EMBL: 229448; NID: 9443858; PIDN: CAA82586.1; PID: 9443859
A; Experimental source: genocitye 1, N3
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 943859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 943859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; NID: 9443859; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229448; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229488; PIDN: CAA82586.1; PID: 9443859
A; Cross-references: EMBL: 229488; PID: 9443859
A; Cross-references: EMBL: 229488; PID: 9443859
A; Cross-references: EMBL: 229488; PID: 9443859
A; Cross-r
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LIPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 by
                                                                                                                                                                                                                                                                                                                                              Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
                                                   A:Experimental source: genotype 2, N3
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-115/Product: core protein #status predicted KM1-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
92.9%; Score 143; DB 2;
Best Local Similarity 92.9%; Pred. No. 6.6e-13;
Matches 26; Conservative 1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                        Score 143; DB 2;
Pred. No. 6.6e-13;
1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 92.9%;
Matches 26; Conservative
A; Cross-references: EMBL: 229457
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 11:05:41 ; Search time 4.90909 Seconds
268.226 Million cell updates/sec
```

Title: US-09-491-146A-28
Perfect score: 154
Sequence: 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28

Sequence: 1 PKPQRKTKRNAHRRPQDVKFPC Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched: 127863 seqs, 47026705 residues tal number of hits satisfying chosen parameters:

127863

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description		Q00269 h genome po	hepatitis	Q01404 hepatitis c	001403 hepatitis c				h genome	h genome	h genome	၎	ч	Ч	herpes si	homo sa	nicot	fugu		spina		P03416 murine coro	P18448 murine coro	m	Q28193 bos taurus	P21573 xenopus lae	P18761 mus musculu	Q09459 caenorhabdi	P35824 bacillus ci	P48155 manduca sex	0		51816 homo sa	26285 bos t
		POLG_HCVJT	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	IE63_HSV2H	FURI_HUMAN	RK4_TOBAC	RS7_FUGRU	RS7_HUMAN	RK4_SPIOL	NCAP_CVM3	NCAP_CVMA5	NCAP_CVMS	FURI_MOUSE	FURI_BOVIN	YB1_XENLA	CAH6_MOUSE	YQ38_CAEEL	SLAP_BACCI	RS7_MANSE	RS7_XENLA	RIBB_VIBPA	FMR2_HUMAN	F262_BOVIN
DB		т	-	Н	Н	-	-	Н	-	П	Н	Н	<del>, -1</del>	~	~4	Н	~	М	Н	П	Н	7	Н	Н	-4	Н	Ч	Н		₽	٦	-	٦	٦
Length		3010	513	520	520	737	737	3010	3010	3011	3033	3033	3010	3011	512	794	282	194	194	293	454	454	454	793	797	303	317	1531	1616	190	194	218	_	3
& Query Match		٠	92.9	'n	'n	ä	ä		ď.	ď.	ď	ά.	88.3	7	æ	ď.			32.5	٠	•	32.5	ς.		•	•			÷	٥.	30.5	•	٥.	9.
Score	•	144	143	4	4	4	4	143	4	4	4	143	136	135	59	52	53	20	20	20	20	50	20	20	49	48.5	48	48	48	47	47	47	47	46
Result No.		<b>⊢</b> +	7	3	4	S	9	7	<b>c</b> o	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	59	30	31	32	33

O15530 homo sapie	Q9z2a0 mus musculu	055173 rattus norv	P23377 rattus norv	Q14738 h serine/th	P78317 homo sapien	Q03717 mus musculu	095049 homo sapien	O9r244 mus musculu	028653 o serine/th	P26786 saccharomyc	P48164 saccharomyc
PDPK_HUMAN	PDPK_MOUSE	PDPK_RAT	FURI_RAT	2A5D_HUMAN	RNF4_HUMAN	KCB1_MOUSE	ZO3_HUMAN	TRP2_MOUSE	2A5D_RABIT	RS7A_YEAST	RS7B_YEAST
-	Н	Н		-	,	-	7	٦	٦	Н	-
556	559	559	793	602	190	857	933	1172	586	189	190
σ.	6.6	29.9	29.9	29.5	29.3	29.5	29.5	29.5	28.9	28.6	28.6
29	•									_	
46 29	46	46	46	45.5	45	45	45	45	44.5	44	44

## ALIGNMENTS

```
CELLULAR AMINOPEPTIDAGE.
CAPSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 144; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NKED (GLCNAC. ) (POT 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                        InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IRR007094; RNA_pol_BSvir.
Pfam; PF01543; HCV_capsid; 1
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_core; 1.
Pfam; PF01560; HCV_NS1; 1.
      IPR002521; HCV_COFE.
IPR002519; HCV_COFE.
IPR002519; HCV_NS1.
IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
IPR001499; HCV_NS4.
IPR001499; HCV_NS4.
HCV_capsid
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    93.5%;
92.9%;
                                                                                                                                                                                                                              Pfam; PF00271; hellcase_C; 1
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               326573
                                                                                                                                                                                  Pfam; PF02907; HCV_NS3; 1
Pfam; PF01006; HCV_NS4a;
Pfam; PF01001; HCV_NS4b;
Pfam; PF01506; HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 92.9 ses 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                          729
1006
1615
1862
2013
369
369
1107
11165
1319
196
209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ä
                                                                                                                                                                                                                                                                                                          3D-structure.
INIT_MET
                             InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@liberalp.ch).
                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
GROSD (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HC-J2) (HCV).
Hepatitis C virus (isolate HC-J2) (HCV).
Hepatitis C virus (isolate HC-J2) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genotypes.";
Virology 188:331-341(1992).

Virology 188:301-341(1992).

Vir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTEWIAL).
MATRIX PROTEIN (FOTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
NONSTRUCTURAL PROTEIN NS1 (POTEWIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR00251; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01560; HCV_NS1; 1.
Pr000m; PD186062; HCV_NS1; 1.
Pr01protein; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                              ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10074; BAA00968.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11111;
                                                                      POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
RESULT 2
POLG_HCVJ2
                                                                              HIDD DD BERLEY STATE STA
```

ö

Gaps

; 0

le-12;

Pred. No.

1; Mismatches

```
POLG_HCVHK
Q01403;
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                       Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                            RESULT 4
POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                            Matches
                                                                                                                                                                                                                                                                   HIDD BEALT THE FETTE TO DESCRIPTION OF THE FETTE THE PRINT OF THE PRIN
                                                                                                                                 ò
                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein.

REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR BUVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTUBAL PROTEIN NS1/E2 (POTENTIAL).
BY SIMILARITY.
                                                                          Gaps
                                                                                                                                                                                                                                                                               01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-93019030; PubMed-1383400;
ADE K., Inchauspe G., Fujisawa K.;
ADE K., Inchauspe G., Fujisawa K.;
AGENOMIC characterization and mutation rate of hepatitis C virus
'Genomic characterization and mutation and partitis of non-A, non-B hepatitis in Japan.";
of non-A, non-B hepatitis in Japan.";
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN WAND GLYCOPROTEIN E. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND MINA.
                                                                                                                                                                                                                                                                                                                                                                                             Hepatítis C virus (isolate HCV-476) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                          ;
0
                                     Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                 Score 143; DB 1; Length 51
Pred. No. 2.1e-13;
1; Mismatches 1; Indels
55704 MW; 943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                             520 AA
                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL, D10688; BAA01530.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01542; HCV_core; 1.
Pfam; PP01559; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nonstructural
                                 92.9%;
illarity 92.9%;
Conservative 1
                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234
305
418
424
431
449
513 AA;
                           Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Transmembrane;
                                                                                                                                                                                                                                         POLG_HCVH4
Q01404;
                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                      RESULT 3
POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
S
                                                                                                            ò
                                                                                                                                             8
                                                                                                                                                                                                                                           SERFER
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation formatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAPSID MATRIX PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN B (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. . . ) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
6enome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GF92) (GF95); Envelope glycoprotein E2 (GF68) (GF70) (NS1)] (Fragment)
Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLING-93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
IIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protein.
REMOVED FROM CAFSID PROTEIN C BY THE
                                                                                                                                                        ö
                                                                                     Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Coat protein; Envelope protein;
                                                                                                                                                        Indels
520 520
520 AA; 56499 MW; AAl35246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE
                                                                                        Score 143; DB 1;
Pred. No. 2.1e-13;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                  520 AA
                                                                                                                                                                                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Co:
Transmembrane; Nonstructural
                                                                                        92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10687; BAA01529.1; -. PIR; JQ1925; JQ1925.
                                                                                                                                                     26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31644;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a licensee agreement (see http://www.isb-sib.ch/announce/or send an email to licenseelab-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ/EZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ (POTENTIAL).
                                                                                                                                                                                                                                                                                                                           01-70(1992 (Rel. 23, Created)
01-70(1992 (Rel. 23, Last sequence update)
16-007-1992 (Rel. 40, Last annotation update)
16-007-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capaid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NSI and NS2) (Fragment); Hepatitis C virus (isolate HG-75) (HCV).
Viruses; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENOTYPES. ";
VITOLOGY 188:31-31(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MANA.
                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-92230232: PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                   (POTENTIAL).
                                  . .) (POTENTIAL).
   (POTENTIAL)
                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΒY
                                                                                                     Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Coat protein; Envelope protein;
                                                                                                                                         1; Indels
                                                                    56476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CELLULAR AMINOPEPTIDASE
N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
                                                                                                       Score 143; DB 1;
Pred. No. 2.1e-13;
                                                                                                                                                                                                                                                                                                737 AA
                                                                                                                                         1; Mismatches
                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                            Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Co.
                                                                                                       92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D10075; BAA00969.1; -.
                                                                                            Ouery Match
Best Local Similarity 92.90
Watches 26; Conservative
                                                                                                                                                                                                                                                                                                STANDARD;
424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                POLG_HCVJ5
                                CARBOHYD
NON_TER
SEQUENCE
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INIT_MET
                                                                                                                                                                                                                                                             RESULT 5
POLG_HCVJ5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                              385
 FT
                                                                                                                                                                                                         셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY FLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J7) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE=92230232; PubMed=1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                             Length 737;
                                                                                                                                                                                                                                                   3AF699D82AD501B1 CRC64;
                                                                                                                                                                                                                                                                           Score 143; DB 1; L. Pred. No. 3.1e-13; 1; Mismatches 1;
                                                                                       (GLCNAC.
                                                                                                                                              (GLCNAC.
                                         (GLCNAC
                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                      (GLCNAC
                                                           (GLCNAC
                                                                         (GLCNAC
                                                                                                                  (GLCNAC
                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                          737 A.A.
                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                           N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NSI.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                   Œ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                             92.98;
                                                                                                                                                                                                                                                                                            95.98;
                                                                                                                                                                                                                                                 81207
                                                                                                                                                                                                                                                                             Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                   737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID-11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVJ7
P27961;
                                                                                    CARBOHYD
                                                                                                                                                                                                                                                 SEQUENCE
                                                          CARBOHYD
                                                                                                                                               CARBOHYD
                                                                                                                                                            CARBOHYD
                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 22, Last sequence update)
61-AUG-1992 (Rel. 23, Last sequence update)
62-003 (Rel. 42, Last annotation update)
62-003 (Rel. 42, Last annotation update)
63-003 (GPR0) (GPC) (GPR18) (GPR28) (GPR38) (GPR0) (Rel. 42, Last annotation Electrical Protein Rel (GPR0) (RR1); Protease/helicase NS3 (PP0) (Hepacivirin NS2 (P21)
62-3-4.22-7); Protease/helicase NS3 (PP0) (Hepacivirin NS4 (P22))
63-4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RRA-directed RNA polymerase) (EC 2.7.7.48)].
63-003 (Rel. 42, Rel. 42, Rel. 43)
64-003 (Rel. 42, Rel. 43)
65-003 (Rel. 43, Rel. 43)
65-003 (Rel. 43, Rel. 43)
65-003 (Rel. 43, Rel. 43, Rel. 43, Rel. 43)
65-003 (Rel. 43, Rel. 43, Rel. 43, Rel. 43, Rel. 43)
65-003 (Rel. 43, Rel. 43, Rel.
                                                                                                                                                                                                                                            CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDILINE-91140689; PubMed-1847440;
MEDILINE-91140689; PubMed-1847440;
Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,
Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                      THE
                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1487-1500.
MEDLINE-9623224; PubMed-8647104;
MCDLINE-9623224; PubMed-8647104;

*Norowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;

*Nor-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                      Β¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143; DB 1; Length 737;
Pred. No. 3.1e-13;
1; Mismatches 1; Indels
                                                                                                                                                   Nonstructural protein.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Probom: PF01560; HCV_NS1; 1.
Probom: PD186062; HCV_NS1; 1.
POlyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81691 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 Similarity 92.9
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MCBI_TaxID-11105;
                                                                                                                                                       Transmembrane;
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P26663:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
POLG_HCVBK
                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
      NAME OF THE STATE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               . E E E E E E E E E E E E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 825E
```

```
This. SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                MEDLINE-9827846; PubMed-9568891;

X MEDLINE-9827846; PubMed-9568891;

A StelnKuehler C., Tomel L., Gardana V.,

StelnKuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
virus: a 2.2-A resolution structure in a hexagonal crystal form.";

Protein Sci. 7:837-847(1998).

- 1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAX A ROLE IN THE VIRAL RNA REPLICATION.

C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral postition, Cys or Thr in pl and Ser or Ala in Pl'.

C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                  [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                           MEDLINE-97015088; Pubmed-8865916; Love R.A., Parge H.E., Wickersham J.A., Hostcmsky Z., Habuka N., Moomaw E.W., Adachi T., Hostcmska Z.; The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site."; Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                          -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                K-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002868; HCV_NS5a.
IPR002166; HCV_RdRP.
IPR007095; RNA_DOl_DS_PS.
IPR007094; RNA_DOl_DS_VIr.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-APR-98.
15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1A1Q; 25-MAR-98
1JXP; 14-JAN-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002522;
IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20-APR-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-NOV-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-JUN-00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                $29.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1GX6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IGX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10UV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BOHM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB;
```

PF01543; HCV\_capsid; 1. PF01542; HCV\_core; 1.

nterPro

etam;

HCV\_NS4a; HCV\_NS3

PF01006;

Gaps

ô

Length 3010;

Score 143; DB 1; Pred. No. 1.4e-12; 1; Mismatches 1

```
1187 1188
1189 1197
1198 1202
11203 1204
1680 1688
3010 AA; 327189 MW; F8422D5ECCFDFD9C CRC64;
                                                Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                          POLG_HCVJA
P26662;
                                                                                                                                                                                                                         Hepacivirus.
                            STRAND
  TURN
                                                                                                              RESULT 8
POLG_HCVJA
                      TURN
  FT
FT
FT
SO
                                                                                        a
                                                                                                                           à
                                                                         CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
POTENTIAL.
                                 Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                             REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                (POTENTIAL).
                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                       POTENTIAL
                                                                    CELLULAR AMINOPEPTIDASE
                                                                                                                                            CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                             N-LINKED (GLCNAC.
                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                 LINKED
                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                     LINKED
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00999; Viral_RGRP; 1.
ProCom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                            645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                              1097
1097
1107
1120
1120
1133
                                                                                                                                                                                                                                                                                                                     035
                                                                                                                                                                                                                                                                                   2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                     1031
1039
10039
10059
10068
10090
11009
11108
11108
11129
11129
                                                       3D-structure.
INIT_MET
                                                                                                                                                         ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                     TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                  STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                      TURN
STRAND
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                            TURN
                                                                                                                                                                                                                                                                                                                     STRAND
                                                                          CHAIN
CHAIN
CHAIN
                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                         CHAIN
                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                            HELIX
                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                         TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                           01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-50G-1992 (Rel. 23, Last sequence update)
01-60G-1992 (Rel. 4), Last annotation update)
Genome polyprotein [Contains notation of Core protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein El (GP83) (GP35); Envelope glycoprotein El (GP83) (GP32) (GP35); Envelope glycoprotein El (GP83) (GP32) (GP32) (GP32); Envelope glycoprotein El (GP83) (GP32) (GP33) (GP32); Nonstructural protein NS4 (P4); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P56); Nonstructural protein NS5 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate Japanese) (HCV)
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [RNA][N].
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
"Molecular Cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                      PRT; 3010 AA.
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D90208; BAA14233.1; -.
                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; A39253; GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11116;
```

**5**8

```
PKPQRKTKRNAHRRPQDVKFPGGGQIVG
                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACTIVATION OF NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RNA (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                        POLG_HCVH
P27958:
                           -
                                                                           ນ
                                                                                                                                                     RESULT 9
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ÷
                           ò
                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 143; DB 1; Length 3010; Pred, No. 1.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                       INTERPOSE | 1PR002521; HCV_core.

INTERPOS | 1PR002521; HCV_core.

INTERPOS | 1PR002531; HCV_core.

INTERPOS | 1PR002531; HCV_NS.1.

INTERPOS | 1PR0004109; HCV_NS.3.

INTERPOS | 1PR000445; HCV_NS.4.

INTERPOS | 1PR001409; HCV_NS.4.

INTERPOS | 1PR001409; HCV_NS.4.

INTERPOS | 1PR001409; HCV_NS.4.

INTERPOS | 1PR001409; HCV_RGRP.

INTERPOS | 1PR001409; HGI_COS.PS.

INTERPOS | 1PR007094; RNA_POL_PS.PS.

INTERPOS | 1PR007094; RNA_POL_PS.PS.

INTERPOS | 1PR007094; RNA_POL_PS.PS.

INTERPOS | 1PR007094; RNA_POL_PS.PS.

INTERPOS | 1PR007094; RNA_POL_PS.PS.
                                                                                                       HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; Pr01560; HCV_NS1; I.Pfam; Pr01560; HCV_NS2; I.Pfam; Pr01906; HCV_NS3; I.Pfam; Pr01006; HCV_NS4b; I.Pfam; Pr01001; HCV_NS4b; I.Pfam; Pr001506; HCV_NS5a; I.Pfam; Pr00998; Viral_RGRP; I.Pr0Dom; PD186062; HCV_NS1; I.Pr0Dom; PD186062; HCV_NS1; I.Pr0Dom; P0186062; P0186062; I.Pr0Dom; P0186062; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.98;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01542; HCV_core;
PF01539; HCV_env;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2013
3010
369
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1165
                                                                         InterPro; IPR001410;
InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001650;
                                                                                                                                                                                                   InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                            InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                            InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3010 AA;
                     MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                    interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
CORRECT TO THE FET THE
```

ö

Gaps

ö

1; Indels

1; Mismatches

Conservative

Best Local Similarity Matches 26; Conserv

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              s its content is in no way
Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on it use by non-profit institutions as long as its content is in no wa modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 22, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
67-00 (GP60) (GR1); Protein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP60) (GR1); Protein E7; Nonstructural protein NS2 (P21)
(EC 3.4.99.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.99.-); Protease/helicase NS3 (P70) (Hepacivirin)
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66); P070) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepacitis C virus (Laolate H) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; Pubmed-9187654;
Yao N., Hesson T., Cable M., Hong 2., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
Structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3 WOUTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4A-NS4B, AND NS5A-NS5B, FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
ESSENTIAL ROLE IN THE VIRUS REPLICATION.
CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
CATALYTIC ACTIVITY: N nucleoside triphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LOJAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9815431; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY JUPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Prince A.M.; define the human prototype strain H of hepatitis C figenomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                           PRT; 3011 AA
                    32
```

```
. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 143; DB 1; L
Pred. No. 1.4e-12;
1; Mismatches 1;
  N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
  1636
1652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1584 159
1598 159
1606 161
1614 161
1622 162
1635 163
1640 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
POLG_HCVJ6
ID POLG_HCVJ6
CARBOHYD
STRAND
TURN
HELLIX
TURN
HELLIX
TURN
STRAND
STRAND
STRAND
STRAND
HELLIX
TURN
STRAND
HELLIX
TURN
STRAND
HELLIX
TURN
STRAND
STRAND
HELLIX
TURN
STRAND
STRAND
HELLIX
HELLIX
HELLIX
HELLIX
HELLIX
HELLIX
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
HELLIX
HELLIX
HELLIX
TURN
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
HELIX
SEQUENCE
  (See http://www.isb-sib.ch/announce/
                       EMBL: M67463 AAA45534.1; -.

DR PIB: AA6814; GNWYCH.

DR MEROPS; 529.001; -.

DR INCEPTO: IPR0015212; HCV_core.

DR INCEPTO: IPR0015213; HCV_core.

DR INCEPTO: IPR0015213; HCV_core.

DR INCEPTO: IPR0015219; HCV_core.

DR INCEPTO: IPR0015219; HCV_core.

DR INCEPTO: IPR001519; HCV_NS3.

DR INCEPTO: IPR00156; HCV_NS3.

DR INCEPTO: IPR00156; HCV_NS3.

INCEPTO: IPR00156; HCV_NS3.

INCEPTO: IPR00156; HCV_CORP.

DR Ffam; PF0154; HCV_CORP.

DR Ffam; PF0154; HCV_CORP.

DR Ffam; PF0154; HCV_NS3; I.

DR Ffam; PF0155; HCV_NS3; I.

DR Ffam; PF01001; HCV_NS3; I.

DR Ffam; PF01005; HCV_NS4; I.

DR Ffam; PF010098; HCV_NS4; I.

DR Ffam
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR ANINOPEPTIDASE.
CARSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5B.
PROTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
ities requires a license agreement (St
send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
CARBOHYD
  entities
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
```

Gaps

```
D POLG_HCVJ8 STANDARD; PRT; 3033 AA.

POLG_HCVJ8 STANDARD; PRT; 3033 AA.

10-26661;

TO 1-40G-1992 (Rel. 23, Last sequence update)

TO 1-40G-1992 (Rel. 23, Last sequence update)

TO 1-40G-1992 (Rel. 41, Last annotation update)

TO 1-40G-1992 (Rel. 41, Least annot
                                                                                                                                 TOTOLINIA STATEMENT TANDER TO THE STATEMENT ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143; DB 1;
Pred. No. 1.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 1.46
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LSITE
BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJ8
          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-9204440; PubMed-1658196;
A MEDLINE-9204440; PubMed-1658196;
A Machida A., Miyakawa Y., Mayumi M.,
Amachida A., Miyakawa Y., Mayumi M.,
Amachida A., Miyakawa Y., Mayumi M.,
Amachida Sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
HYDROPHOBIC, SUGGESTING A POSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROFE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROMMONLY WITH ASP OF Glu in the viral position, Cys or Thr in Pl and Ser or Ala in Pl'.
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                            01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-60G-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS48 (P77); Nonstructural protein NS58 (P65) (P77); Nonstructural protein NS58 (P77); Nonstructural P77 (P77); NS58 (P77); NS58 (P77); NS58 (P77); NS58 (P77); NS58 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ro; IPR007095; RNA_pol_DS_PS.
ro; IPR007094; RNA_pol_PSvir.
PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_core.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR004109;
InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P27958; 1HE
MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       - # \frac{1}{2} \cdot \c
```

POTENTIAL)

ö

Gaps

ö

POTENTIAL

POTENTIAL

SIMILARITY). SIMILARITY). SIMILARITY).

(BY (BY (BY

(POTENTIAL) POTENTIAL) (POTENTIAL) ( POTENTIAL) (POTENTIAL) POTENTIAL)

(GLCNAC.

(GLCNAC

(POTENTIAL)

(GLCNAC GLCNAC (GLCNAC (GLCNAC (POTENTIAL) (POTENTIAL)

GLCNAC (GLCNAC (GLCNAC

(POTENTIAL)

(POTENTIAL)

(POTENTIAL)

(GLCNAC

(POTENTIAL)

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein (Contains notation update)
66-00me polyprotein (Contains notation (P22);
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein El
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein El
67-00me polyprotein El (GP32) (GP35); Envelope glycoprotein El
68-3-4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P67); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)}.
Hepatitis C virus (Isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISIS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1A173E7E3381FD1A CRC64;
                                           CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 143; DB 1;
Pred. No. 1.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 1.4e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-92230206; PubMed-1314449;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2811
330177 W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match 92.9%;
Local Similarity 92.9%;
les 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   649
1091
2038
2359
2359
2311
2811
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (RNA) (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVTW
P29846;
                                                                                          ACT_SITE
NP_BIND
         CHAIN
TRANSMEM
                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                          CARBOHYD
                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation retropped Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
Transmembrane; Nonstructural process.
Transmembrane; Nonstructural proversing CELLULAR AMINOPEPTIDASE.
Transmembrane; Nonstructural AMINOPEPTIDASE.
Transmembrane; Nonstructural PROTEIN (POTENTIAL).
Transmembrane; Nonstructural PROTEIN (POTENTIAL).
Transmembrane; Nonstructural PROTEIN NS1 (POTENTIAL).
Transmembrane; Nonstructural PROTEIN NS1 (POTENTIAL).
Transmembrane; Nonstructural PROTEIN NS4 (POTENTIAL).
                                                                             *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *, virology 188:331-341(1992)

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
                    MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                         SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY 529.
                                                                                                                                                                                                                                              precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1', CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Interpro; IPR002518; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
Interpro; IPR00145; HCV_NS4a.
Interpro; IPR001490; HCV_NS4b.
Interpro; IPR002868; HCV_NS5a.
Interpro; IPR007096; HCV_MSA.
Interpro; IPR007095; RNA_POL_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV_capsid,
InterPro; IPR002511; HCV_core,
InterPro; IPR002519; HCV_env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
Probom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF01543; HCV_caps1d;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01542; HCV_core;
PF01539; HCV_env; 1
PF01560; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF02907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
DORKEN KEN CORRECT COR
```

ö

Gaps

ö

Length 3033; 1; Indels K

```
EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DR
                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MAJOR ENVELOP PROTEIN (ALLA).
MAJOR ENVELOPE PROTEIN SILEZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSILEZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Primmi Pr01342; HCV_capsic, 1.

R Pfam; PF01542; HCV_cnv; 1.

R Pfam; PF01542; HCV_cnv; 1.

R Pfam; PF01560; HCV_NS2; 1.

R Pfam; PF01006; HCV_NS2; 1.

R Pfam; PF01006; HCV_NS4; 1.

R Pfam; PF01001; HCV_NS4; 1.

R Pfam; PF01001; HCV_NS5; 1.

R Pfam; PF00271; HcV_NS5; 1.

R Pfam; PF00998; Viral_RGRP; 1.

R Probom; P0186062; HCV_NS1; 1.

R Pfam; PF00998; Viral_RGRP; 1.

R Probom; P0186062; HCV_NS1; 1.

R Probom; P0186062; HCV_NS1; 1.

R Pfam; PF00998; Viral_RGRP; 1.

R Probom; P0186062; HCV_NS1; 1.

R Pfam; PF00998; Viral_RGRP; 1.

R Probom; P0186062; HCV_NS1; 1.

R Pfam; PF00998; Viral_RGRP; 1.

R Pfam; PF01098; Viral_RGRP; 1.

R Pfam; PF01099; Viral_RGRP; 1.

R Pf01099; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC...) (POTENTIAL).
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INTERPROPTION DEAD.

INTERPROPTION DEAD.

INTERPROPTION THROUGHOUS THROUGH THROU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
PDB; 1N64; 25.FEB-03.
PDB; 1N83; 08-APF-98.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116
192
384
730
1007
1616
1863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        THE REAL PROPERTY OF THE REAL
```

HHYD 430 439  HHYD 532 532  HHYD 540 544  HHYD 540 546  HHYD 540 546  HHYD 526 556  526  HHYD 526 526  HHYD 527 520  HHYD 2041 2041  2041 2041  2041 2041  2041 2041  2042 2042  2042 2042  2043 2043  2043 2043  2044 2043  2041 2044  2041 2044  2041 2044  2041 2044  2041 2044  2041 2044  2042 2043  2043 2044  2044 2044  2049 2044  2049 2049	N-LINKED (GLCNAC) (POTENTIAL).	Score 136; DB 1; Length 3010; Pred. No. 1.5e-11; 1; Mismatches 1; Indels 0; Gaps 0	PPGGQIVG 28           PPGGQIVG 32	PRT: 3011 AA.	uence updat	(Core prote	elope glycoprotein E2 ral protein NS2 (P21) (Hepacivirin)	tural protein NS4A (P4); Nonstructural protein 1 protein NS5A (P56); Nonstructural protein rected RNA polymerase) (EC 2.7.7.48)	) (HCV). nd viruses, no DNA stage; F]		18704	J.H., Berger K., Lee C., Dong C.,	of the hepatit	2451-2455(1991). NS2A, NS2B, NS4A AND N	THE VIRAL RNA REPLICATION.	of four peptide y with Asp or Glu	r in Pl and Ser or Ala in Pl'. N nucleoside triphosphate = N diphosphate +	IS A NUCLEOCAPSI	CONSISTS OF TWO PROTEINS:	-	. e	There are no re			, , , , , , , , , , , , , , , , , , , ,
	430 448 532 532 540 576 623 623 2204 2204 2220 2288 33	h Similarity 92.6% 25; Conservative			(Rel. 23, Creat (Rel. 23, Last	រួយដួ	be glyco (GP70) 1.22);	1.21.98) 27); NO 366) (P7	18 C virus (1sola:	Hepacivirus. NCBI_TaxID=11104;	I.) EQUENCE FROM N.A. FPDLINE-91172826: PubMed-18	Choo QL., Richman K.H., H	O.W., Kuo G., organization	Natl. Acad. Sci.	NS3 AND NS5 MAY PI	- CATALYTIC precursor	Cys or Th ACTIVITY:	(RNA)(N).	LIPOPROTEIN ENVELOPE. T PROTEIN M AND GLYCOPROT		his SWISS-PROT entry is co etween the Swiss Institut	he European Bioinformatics see by non-profit instit	ified an	send an	TI W62321 . LECT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   bolan A.;
submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
-- FUNCTION: INVOLVED IN THE DOWN-REGULATION OF VIRAL IMMEDIATE-EARLY
GENES AND ACTS IN COMBINATION WITH ICPO AND ICP4 AS AN ACTIVATOR
OF LATE GENES (BY SIMILARITY).
-- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
HSV-2 UL54, EHV-1 S, V2V 4, EBV BMLFI, HCMV UL69, AND HYS-1 57.
                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-92113549; PubMed-1662697; McGeoch D.J., Cunningham C., McIntyre G., Dolan A.; Conningham C., McIntyre G., Dolan A.; Conningham C., McIntyre G., Dolan A.; Conningham C., McIntyre G., Dolan C., G., McGeoch D.J., Golining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2."; J., Gen. Virol. 72:3057-3075(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Early protein; Transcription regulation; Activator; DNA-binding.
SEQUENCE 512 AA; 54958 MW; 459651470A503BA7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FURI_HUMAN STANDARD; PRT; 794 AA.
P09958; Q14336;
O1-MAR-1989 (Rel. 10, Created)
O1-MAR-1990 (Rel. 14, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 3.4.21.75) (Paired basic amino acid residue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ..
                                                                                                        ö
                                                         Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                38.3%; Score 59; DB 1; Length 512; 40.6%; Pred. No. 0.47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12; Indels
       327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Herpes simplex virus (type 2 / strain HG52).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
NCBI_TaxID-10315;
                                                 Query Match 87.7%; Score 135; DB 1; Lk Best Local Similarity 85.7%; Pred. No. 2.1e-11; Matches 24; Conservative 2; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                           01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Transcriptional regulator IE63 (VMW63) (ICP27).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNAH----RRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                   512 AA
                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3; Mismatches
                                                                                                                                                                        PKPQKKNKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10471; BAA01269.1; -. EMBL; 286099; CAB06702.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 40.6
Matches 13; Conservative
                                                                                                                                                                                                                                                                                                                   STANDARD;
         3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; JQ1498; WMBEXA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                              IE63_HSV2H
P28276;
       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                  RESULT 14
IE63_HSV2H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
FURI_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8444B
                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                            ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN B (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI/F2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/F2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein; Clycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                           Interpro; IPR002166; HCV_RdRP.
Interpro; IPR001650; Hallaase_C.
Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
                                                                                                                                           HCV_core.
                                                                                                                                                                                     InterPro; IPR002519; HCV_env.
InterPro; IPR002511; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4a.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06; HCV_NS4a; 1.
01; HCV_NS4b; 1.
06; HCV_NS5a; 1.
71; helicase_C; 1.
98; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01543; HCV_capsid; Pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_env; 1.
                                                                                                                    InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
InterPro; IPR002521; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProDom; PD186062; HCV_SMART; SM00487; DEXDC;
A39166; GNWVC3.
1A1V; 16-FEB-99.
1HEI; 25-NOV-98.
                                                                     MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11616
11863
2014
11083
11083
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01539;
PF01560;
PF01538;
PF02907;
PF01006;
PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

LANGURAR MANAGERAR AND CORRECT CORRECT

ï

EMBL outstation

and the

the Swiss Institute of Bioinformatics

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREBEQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPLASMIC RETICULUM (ER). SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK (TGN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN.
                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 296-794 FROM N.A. MEDLINE-87053858; PubMed-3023061; MEDLINE-87053858; PubMed-3023061; Roebroek A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.; "Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDINE-94291619: PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
"Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilish-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                  "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.
DOMAIN: CONTAINS A CYTOPLANIC DOMAIN RESPONSTBLE FOR ITS TGN
LOCALIZATION AND RECYCLING FROM THE CELL SURRAGE.
FTM: THE PROPEPTIDE IS AUTOCATALYTICALLY REMOVED THROUGH AN
INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM
(ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD IT
TO THE ACTIVATION OF FURIN
                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
                                                                                                                                                          "Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2."; Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                             Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Klefer M.C.,
                                                                                                          TISSUE-Blood;
MEDLINE-9017502; PubMed-2408021;
Van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D.,
Dorssers L.C.J., van de Ven W.J.M.;
cleaving enzyme) (PACE) (Dibasic processing enzyme). FURIN OR PACE OR FUR.
                                                                                                                                                                                                                                                                                                                                                                                                                                   3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDLINE=91321735; Pubmed=1713771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COFACTOR: CALCIUM-DEPENDENT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       respective precursors.
                           Homo sapiens (Human).
                                                                                              SEQUENCE FROM N.A.
                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                            Brake A.J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              788888888888888888888888888
```

```
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SYSTEM (BY SIMILARITY).
SYSTEM (BY SIMILARITY).
SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Serine protease; Transmembrane; Glycoprotein; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC. . .) (POTENTIAL CLEAVAGE (SECOND AUTO-).
CLEAVAGE (FIRST AUTO-).
CELL ATTACHMENT SITE (POTENTIAL).
CELL SURFACE SIGNAL.
TRANS GOLGI NETWORK SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 794;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                        GO, GO: 0004276; F: furin activity; TAS.
GO; GO: 0007267; P: cell-cell signaling: TAS.
GO; GO: 000508; P: protelolysis and peptidolysis; TAS.
INTERPRO; IPR006212; Furin_repeat.
INTERPRO; IPR006209; Pedmain.
INTERPRO; IPR00209; Peptidase_S8.
Pfam; PF01481; P_Proprotein; PARTIAL.
Pfam; PF01082; Peptidase_S8; 1.
PRINTS; PR00723; SUBTILISIN.
PRODM; PD000717; P_Gomain; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10C44DD5892EF35D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .,
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 55; DB 1;
Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARGE RELAY S
CHARGE RELAY S
CHARGE RELAY S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                         GO:0005794; C:Golgi apparatus; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CYS-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS00136; SUBTILASE_ASP; PROSITE; PS00137; SUBTILASE_HIS; PROSITE; PS00138; SUBTILASE_SER;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :|:|||: :: | | |||
102 KRRTKRDVYQEPTDPKFP 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 QRKTKRNAHRRPQDVKFP 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86678 MW;
                                                                                                                     EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA0660S.1; -. PIR; A39552; KXBMF. HSSP; Q99405; 1MPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35.7%;
50.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                         MEROPS; S08.071; -. Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     794 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Symogen; Calcium
                                                                                                                                                                                                                                                                    MIM; 136950;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SITE
```

Search completed: August 7, 2003, 11:20:01 Job time : 4.90909 secs

```
MEDITIE—56305314; PubMed-8712927;
Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba K.-I., Wu R.-R.,
Ohno T., Mizokami M., Saleh M.G., Orito E., Ohba K.-I., Wu R.-R.,
Koide T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R.;
Wosefulness and limitation of phylogenetic analysis for hepatitis c
virus core region: application to isolates from Egiptian and Yemeni
patients.;
-1- SUBUNT: THE VIRION OF THIS VIRIS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 1
105 105
105 AA; 11870 MW; C7BA40B284025A49 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D82034; BAA11519.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER
NON_TER
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              000000000000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hepatitis
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P90364 hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hepatitis
                                                                                                                                       August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     068575 | 091au0 | 091au0 | 091au0 | 091591 | 0068311 | 0068306 | 0081328 | 0081328 | 0081328 | 0081327 | 0081326 | 0081327 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0081326 | 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                              830525
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              stal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                              830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                               - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q68575
Q91AU0
Q81989
Q81541
Q68311
Q68306
Q68309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8JYR7
Q8JYR6
Q68712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QBJYRB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q8JYR5
                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΩI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_organelle:*
                                                                                                                                                                                                                               US-09-491-146A-28
154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SPTREMBL 23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_fundi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191
3010
3010
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_mhc: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
                                                                                                                                                                                                                                         Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                               OM protein
                                                                                                                                                                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                 Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database
                                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
```

17	-	ď	74	75	068682	068682	hepatitis c	
18	*	~;	74	12	068708	068708	hepatitis	
19	4	ď	74	12	068679	068679	hepatitis c	
20	**	ď.	74	12	Q68683	068683	hepatitis	
21	*	ď	74	12	Q68706	068706	hepatitis	
22	*	ď	74	12	068684	Q68684	hepatitis	
23	7	ď	74	12	968707	068707	hepatitis c	
24	**	ď	74	12	068685	068685	hepatitis	
25	4	ď	16	12	Q8JYS0	Q81ys0	hepatitis	
26	*	ď	78	12	Q8JYS1	08jys1	hepatitis	
27	<b>**</b>	ď	83	12	081264	081264	hepatitis	
28	143	95.8	100	12	Q9QT61	09qt61	hepatitis c	
59	₹*	ς.	100	12	Q8QP70	08qp70	hepatitis	
30	-	ď	100	12	Q8QP54	Q8qp54	hepatitis	
31	-	ď	100	12	Q9QT57	09qt57	hepatitis	
32	₹*	ď.	100	12	Q8QP52	Q8qp52	hepatitis	
33	4	ď	100	12	Q8QP88	08qp88	hepatitis	
34	~	'n	100	12	Q9QT56	09qt56	hepatitis	
35	4	ς.	100	12	Q8QP90	08db80	hepatitis	
36	4	'n.	.100	12	Q8QP64	08qp64	hepatitis	
37	4	ä	100	12	Q8QP55	Q8qp55	hepatitis	
38	4	ď	100	13	Q8QP60	09db60	hepatitis	
39	4	ς.	100	12	Q8QP57	08qp57	hepatitis	
40	₹	'n	100	12	Q9QT58	09qt58	hepatitis	
41	ゼ	ς.	100	12	Q8QP63	08qp63	hepatitis	
42	4	ä	100	12	Q8QP61	Q8qp61	hepatitis	
43	4	ä	100	12	Q9QT52	09qt52	titis	
44	4	ď	100	12	Q9QT54		itis	
	4	ä	100	17	Q8QP59	08qp59	itis	
					ALIGNMENTS			
RESULT 1								

```
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                          068575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091AU0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   091AU0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q91AU0
                                                                                                                                                                                                                                                                                                                                                                          DDT THE SECOND DESTRICT OF THE SECOND DESTRIC
                                                                                                                                                                                                                   a
                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; saRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepcivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (FEB-1995) to the EWBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C MND MRNM (BY SIMILARITY).
EMBL; D49465; BAA08439.1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus type 2.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-94351179; PubMed-7520922;
Nakazawa T., Kato N., Ohkoshi S., Shibuya A., Shimotohno K.;
J. Hepatol. 20:63-629(1994).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
InterPro: IPRO02522; HCV_capsid.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ohno T., \texttt{Mizokami} \ \texttt{M.}: "Determination of nine genotypes of hepatitis C virus using PCR
                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 144; DB 12; Length 106;
Pred. No. 6.9e-14;
1; Mismatches 1; Indels C
Score 144; DB 12; Length 105;
Pred. No. 6.8e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-2002 (TrEMBLrel. 22, Last annotation update) (isolate USA8) genomic RNA (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106
12001 MW; 25D0D5414B3EA9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D945CB60661797E3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09PXN3;
01-MAY-2000 (TIEMBLrel. 13, Created)
01-MAY-2000 (TIEMBLrel. 13, Last sequence update)
01-CCT-2002 (TIEMBLrel. 22, Last annotation update)
EZ/NSI protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                       106 AA
                                                                                                                       1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro, IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125 AA; 14220 MW;
93.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ch 93.5%;
1 Similarity 92.9%;
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01,
                            Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        106 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=40271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-USA8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
SEQUENCE 1
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mehtod.
                                                                                                                                                                                                                                                                                                                                    081831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9PXN3
                                                                                                                                                                                                                                                                     RESULT 2
081831
1D 0818
AC 0818
DT 01-N
DT 01-O
DT 01-O
DE (150
OS HEPA
OC VITU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TESULT 3
                                                                                                                       õ
                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPART REPART OF THE PART OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lim S.P., Khu Y.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.; "Identification and molecular characterization of the complete genome of a Singapore isolate of hepatitis C virus: sequence comparison with other strains and phylogenetic analysis."; virus Genes 23:89-95(2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
20840 MW; 4AAE63444D8329E2 CRC64;
                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
Score 144; DB 12;
Pred. No. 8.2e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93.5%; Score 144; DB 12;
92.9%; Pred. No. 1.3e-13;
live 1; Mismatches 1;
                                                                                                                                                                                                                   191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3010 AA.
                                                                                             5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                             1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-HCV-S1;
MEDLINE-21440119; Pubmed-11556407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
93.5%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01543; HCV_capsid;
Pfam; PF01542; HCV_core; 1.
                    Best_Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 92.9
                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE 191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                 [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
```

```
single Japanese carrier in Nagasaki prefecture and genome analysis of El and E2/NS1 envelope glycoprotein regions.";
Jpn. J. Trop. Med. Hyg. 22:169-177(1994).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; D63857; BAA09919.1; --
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.; "Molecular cloning of hepatitis C virus genome from a single Japanese Carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93.5%; Score 144; DB 12; Length 3010; 92.9%; Pred. No. 2.4e-12; ive 1; Mismatches 1; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hydrolase, Nonstructural protein, Polyprotein,
RNA-directed RNA polymerase, Transferase, Transmembrane.
SEQUENCE 3010 AA, 327503 MW; C7BDB38169D6E3CF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      081541;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                            IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                             HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                          HCV_NS5a.
HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00271; helicase_C; 1
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                 PF01543; HCV_capsid; 1
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4b;
HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                             InterPro; IPR002522;
                                                                                                                                                                                                                                     IPR002531;
                                                                                                                                                                                                                                                      IPR002518;
                                                                                                                                                                                                                                                                                         IPR000745:
                                                                                                                                                                                                                                                                                                         IPR001490;
                                                                                                                                                                                                                                                                                                                                                             IPR001650;
                                                                                                                                                                                                                                                                      IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                           InterPro; IPRO0
                                                                                                                                                                                                                                                                                                                                                                                nterPro:
                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                               interPro;
                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                            interPro;
                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                  nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         081541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         281541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC; 1.
SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VARAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase, Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zheng W.Z.;
"Genotype identification of hepatitis c virus (HCV) isolated from a
                                                 Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBMNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93.5%; Score 144; DB 12; Length 3010; 92.9%; Pred. No. 2.4e-12; 1ve 1; Mismatches 1; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q81989;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAN-2003 (TrEMBLrel. 23, Last annotation update)
El and E2/NS1 envelope glycoprotein (Genome polyprotein).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRNAHRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                          EMBL; AF356827; AAL00900.1; -.
InterPro; IPR000345; Cytc_heme_bind.
                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_Vir.
Pfan: PF01543; HCV_capsid; 1.
Pfan: PF01542; HCV_core; 1.
                                                                                                                                                                                            IPR002522; HCV_capsid.
IPR002521; HCV_core.
IPR002519; HCV_env.
                                                                                                                                                                                                                                                  IPR002531; HCV_NS1.
IPR004109; HCV_NS3.
IPR004109; HCV_NS3.
IPR001490; HCV_NS4b.
IPR001490; HCV_NS4b.
IPR002868; HCV_NS5b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                               InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1
PF01001; HCV_NS4b; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 92.9
hes 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                  STRAIN-HCV-S1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zheng W.-Y.Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Zheng W.Z.
                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                     Lim S.P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       081989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
Q81989
```

SO WE WAY TO SERVE TO

Gaps

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
NON_TER
SEQUENCE 45
                                                                                                                                                                                                                  Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q68306
Q68306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        068309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                068309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 9
Q68306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SHW BRACCCC
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                     K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polymerase; Transferase; Transmembrane
Virus Res. 23:39-53(1992).

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GIVCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D1135; BAA1894.1; --
HSSP; P26663; 1JXP.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
SYRAIN-HOV-BBB3,
SONGSIVILAI S., Kanistanon D., Kunkitti R.;
Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 144; DB 12; Length 3010; Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE: PS50507; RDRP_POSITIVE: 1.
PROSITE; PS50501; RDRP_VIRAL; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural;
Polyprotein; Rna-directed RNA polymerase; Transferase; Transferase; Transferase; Transfer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        05F0B2102CF9DD9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOY-1996 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 2.46
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NS4A.
                                                                                                                                                                                                                                                                                                                                                             ; IPR000745; HCV_NS4a.
; IPR001490; HCV_NS4b.
; IPR001868; HCV_NS5a.
; IPR00166; HCV_RdRp.
; IPR007095; RNA_pol_DS_PS.
; IPR007094; RNA_pol_PSv1r.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                          IPRO0252; HCV_capsid.
IPRO02521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ) PF02907; HCV_NS3; 1.

PF01006; HCV_NS4; 1.

PF01001; HCV_NS4b; 1.

PF01506; HCV_NS4b; 1.

PF01506; HCY_NS4b; 1.

PF00998; VITAL_RGRP; 1.

PF0186662; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                             HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        93.5%;
92.9%;
                                                                                                                                                                                                                                                                                          HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        326564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_core;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3est Local Similarity 92,9
4atches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_env;
HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SM00487; DEXDc;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01538; HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2014
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                384
730
1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00998
                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                             nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q68311
Q68311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
06831
1D 06831
1D 0683
1D 0683
1D 01-W
1D 01-
     CONTRACTOR OF THE PROPERTY OF THE PARTY OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Songsivilai S., Kanistanon D., Kunkitti R.; "Identification and characterisation of Thai isolates of hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                  4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIKION OF THIS VIROS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; U23749; AAA65056.1; -- InterPro; IPRO0252; HCV_capsid.
                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                        Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 143; DB 12; Length 45; Pred. No. 3.9e-14;
                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                             CC527167096AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CC4C21ED236AAA81 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                    Score 143; DB 12;
Pred. No. 3.9e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Last sequence upd. 01-OCT-2002 (TrEMBLrel. 22, Last annotation u. Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             46 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro: IPR02522; HCV_caps1d. Polyprotein
                                                                                                                                                                                                        45
5015 MW;
                                                                                                                                                                                                                                                                      92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45
5030 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92.9%;
92.9%;
                                                                                                                                                                                                                                                                  Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45
45 AA;
                                                                                                                                                                                                   45 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. STRAIN-HCV-BB5;
```

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
6897 MW;
                                                                                                                                                                                              92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 22, (TrEMBLrel. 22,
                                                                                                                                                                                              Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
es 26; Conserv
                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                  NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-RIG290;
                                                                                                                                                        Polyprotein.
NON_TER
SEQUENCE 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
                                                                                                                                                                                                                                                                                                                            01-OCT-2002
01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2002
01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                    QBJYR5
                                                                                                                                                                                                                                                                                                        Q8JYR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8JYR7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OBJYR7
                                                                                                                                                                                                                                                                                       RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
 .a
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E A E E
                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
       STRAIN-HCV-BB10;
Songsivilal S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                    Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMEL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AF506616; AAM33922.1;
---
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                  DB 12; Length 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 143; DB 12; Length 60;
Pred. No. 5.3e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                     Indels
                                                                                                                                              ABDC4C21ED236AAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60 AA; 6686 MW; B5691CBE8F26F2F1 CRC64;
                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 22, Last sequence update) (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                    1;
                                                                                                                                                                           Pred. No. 4e-14;
; Mismatches
                                                                                                                                                                                                                                                                           60 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 AA
                                                                                                                                                                                                                  1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 143;
                                                                                                                                                                                                        1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                           Genome polyprotein (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Genome polyprotein (Fragment)
                                                                                                                                                                 Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                              5129 MW;
                                                                                                                                                                                                                                                                                               (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                              46 AA;
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-RIG306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER
SEQUENCE 60
                                                                                                                                                                                                                                                                                             01-OCT-2002 (
01-OCT-2002 (
01-MAR-2003 (
                                                                                                                            Polyprotein.
                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QBJYRB;
                                     virus.
                                                                                                                                                                                                                                                                          Q8JYR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q8JYR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
Q8JYR8
                                                                                                                                                                                                                                                       RESULT 11
`.
                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                    S T S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DE DI DI OS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                             K
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Shuston A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbank/DDBJ databases
-!-SIBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED ILIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX (EMBL, AF506620; AAM3396.1;
InterPro; IPRO02522; HCV_capsid.
PFam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                            Score 143; DB 12; Length 61;
Pred. No. 5.4e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 143; DB 12; Length 61; Pred. No. 5.4e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                             61 61
61 AA; 6897 MW; EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EC656DC79E8F26F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Shustov A.V., Gavrilova I.V., Netesov S.V.;
Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Mestern Siberia.";
Submitted (APR-2702) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

REMBL; AF506619; AAM3395.11;
REMBL; AF506619; AAM3395.11;
RPG PROTEIN FF01395.11;
RPG PROTEIN FF0143; HCV_capsid; 1.
RPG PROME FF0143; HCV_capsid; 1.
RPG PROME 62 AA: 7053 MW; EEEC656DC79E8F26 CRC64;
                                                                                                                                                               Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN EVIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL, AF506618; AAM33394.1; -
InterPro; IPR005522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                           Virgorians ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus. NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143; DB 12; Length 62;
Pred. No. 5.5e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                       Polyprotein.
NON TER 61 61
SEQUENCE 61 AA; 6869 MW; EC657F706E8F26F2 CRC64;
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-OMR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 143; DB 12;
Pred. No. 5.4e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: August 7, 2003, 11:19:01 Job time : 26.6364 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                    92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                     Hepatitis C virus.
                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-RIG289;
                                                                                                                                                   STRAIN-RIG288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8JYR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OBJYR6,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
08JYR
AC 08JYR
AC 08JYR
DT 01-0C
DT 01-MA
DE Genom
OS Hepat
OC VHPus
OX NCBL
RR 1ER
RR 5EQUE
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

16, Appl 1, Appli

Sequence 2 Sequence 3 Sequence 3

Sequence ( Sequence 1

Sequence 3 Sequence 3 Sequence 3

Sequence 167,

Sequence 1 Sequence 1 Sequence 1 Sequence 1 Sequence 1 Sequence 1 Sequence 1

```
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOMARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: BNDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAN
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: USA

IP: 30303-1769
COMPUTER READABLE FORM:
MEDIUW TYPE: FIDPPY disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
FILING DATE:
CLASSIFICATION: 435
ATTONREY/AGENY INPORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
RESERENCE/DOCKET WIMBER: 36,714
RESERENCE/DOCKET WIMBER: 36,714
RELEPHONE: 404-818-3790
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTER.STICS:
LEMETH: 28 maine acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Score 154; DB 3;
100.0%; Pred. No. 7.6e-16;
11ve 0; Mismatches 0;
                                     PCT-0895-13552-2

US-07-681-701-16

US-08-290-665A-155

US-08-290-665A-155

US-08-290-665A-157

US-08-290-665A-157

US-08-290-665A-159

US-08-290-665A-161

US-08-290-665A-163

US-08-290-665A-163

US-08-290-665A-163

US-08-290-665A-163

US-08-290-665A-163

US-08-290-665A-163

US-08-290-665A-163

US-08-290-665A-163
                                                                                                                                                                                                                           US-08-290-665A-166
US-08-290-665A-167
                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                       Sequence 28, Application US/08921887; Patent No. 6030771; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein HYPOTHETICAL: NO
 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ANTI-SENSE: NO
FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-921-887-28
                                                                                                                                                                                                                                                                                                                                           US-08-921-887-28
Query Match
2, Appli
1, Appli
1, Appli
9, Appli
23, Appli
3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 28, Appl
Sequence 33, Appl
                                                                                                             (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 191, A Sequence 192, A Sequence 6, Apr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 198,
Sequence 232,
Sequence 14, A
Sequence 8, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence 10,
                                                                                                August 7, 2003, 11:07:41; Search time 10.5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 3
Sequence 2
Sequence 1
Sequence 1
Sequence 9
Sequence 2
Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2
Sequence 9
Sequence 2
Sequence 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                              GenCore version 5.1.6 . . Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-921-887-28
US-08-921-887-33
US-08-921-887-35
US-08-635-886C-191
US-08-635-886C-192
US-08-380-160-6
US-09-020-446-36
US-08-380-160-2
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US95-13660-3
US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-380-160-1
US-07-946-054-9
US-08-033-947-23
US-08-530-550-3
US-08-252-037-26
PCT-US93-09638-9
PCT-US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-836-075A-14
US-08-324-977-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-836-075A-46
                                                                                                                                                                                                                                                                                   tal number of hits satisfying chosen parameters:
                                                                                                                                                                     154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-384-616-8
                                                                                                                                                                                                                                                       328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                     protein search, using sw mode.
                                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                               seq length: 0
seq length: 200000000
                                                                                                                                                        US-09-491-146A-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВВ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match Length
                                                                                                                                                                     Perfect score:
Sequence:
                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score
                                                                                                                                                                                                                                                                                                               Minimum DB
Maximum DB
                                                                       protein
                                                                                                                                                                                                                                                                                                                                                                                                                  Database :
                                                                                                                                                                                                                                                        Searched:
                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
```

ö

Gaps

ö

Length 28; Indels

ò g

```
APPLICANT: LERONGY-ROELS, GEERT
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
TITLE OF INVENTION: UNIVENDED OF HEPATITIS C
TITLE OF INVENTION: UNIVENTION: UNIVENTION: UNIVENTION: UNIVENTION: UNIVENTION: UNIVENTION: UNIVENTION: UNIVERSITY APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1994-10-28
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PATENTIN VERSION 3.1
SEQ ID NOS: 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                             OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
              191 Peachtree Street, N.W., 37th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.5%; Score 144; DB 4;
92.9%; Pred. No. 4e-13;
live 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
93.5%; Score 144; DB 3;
Best Local Similarity 92.9%; Pred. No. 2.1e-14;
Matches 26; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 191, Application US/08635886C Patent No. 6555114 GENERAL INFORMATION:
                                                                  COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DC
SOFTWARE: Patentin Release #1.
                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEPHONE: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; ORGANISM: Hepatitis virus US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                    FILING DATE:
CLASSIFICATION: 435
STREET: LOT
TTY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ANTI-SENSE: N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                              APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 25, Application US/08921887

Patent No. 6030771

GENERAL INFORMATION:

APPLICANT: KHUDYAKOV, YURI E.

APPLICANT: FIELDS, HOWARD A.

TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION

TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; DB 3; Length 28; 5.6e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 96.1%; Score 148; DE Best Local Similarity 96.4%; Pred. No. 5.6e Matches 27; Conservative 0; Mismatches
                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/08/921,887
                                                                                                                                                              Sequence 33, Application US/08921887 Patent No. 6030771 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: USA
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                           US-08-921-887-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-921-887-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGIH:
```

g

```
Length 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE: D.C.

ZIP: 20007-5109

COUNTRY: U.S.A.

ZIP: 70007-5109

COMPUTER READABLE FORM:

MEDIUW TYPE: Floppy disk

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PATENTIN Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/020,846

FILING DATE: U9-FEB-1998

CLASSIFICATION NUMBER: US/09/020,846

FILING APPLICATION TOWNER: P9 -027015

FILING APPLICATION NUMBER: US/09/020,846

FILING APPLICATION NUMBER: US/09/020,846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 143; DB 3;
Pred. No. 3.7e-14;
1; Mismatches 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIWAKUMA, Tomiko
APPLICANT: KASHIWAKUMA, Tomiko
APPLICANT: YAGI, Shintaro
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAMA, Akira
TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPORKTKRNAHRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFENCE/DOCKET NUMBER: 053466/0225
TELECOMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
WPB 28682
                                                                                                                                                                                                                                                                                                                                    ORIGINAL TYPE: N-terminal ORIGINAL SOURCE: ORGANISM: Human Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 36, Application US/09020846 Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET: 3000 K Street, N.W. CITY: Washington
               TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92.9%;
92.9%;
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 3
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                       LENGTH: 34 amino acids
                                                                                                                    INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                          TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                           HYPOTHETICAL: N
ANTI-SENSE: NO
                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: DALBON, Michel
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHERIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                    Sequence 192, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: UNUS
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: IMMUSER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR PLICATION NUMBER: EP 93402718.6
PRIOR PLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTHAMRE: PATENTIN VERSION 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 144; DB 4; Length 450; Pred. No. 4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                 PKPORKTKRNAHRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-WAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Alexandria
STATE: VA
COUNTRY: VA
COUNTRY: 22330
ZIP: 22330
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPALIDLE
COMPUTER: TBM PC STEEM: PC-DOS/MS-DO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 6, Application US/08380160
Patent No. 6235284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BERRIDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Berridge, William P. REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 93.5%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: OLIFF & BER
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                     US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
US-08-380-160-6
                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

á

ö

Gaps

ö

```
APPLICANT: JOLIVET, MICHEL
APPLICANT: DELINI, FRANCOIS
APPLICANT: DELINI, FRANCOIS
APPLICANT: DALBOW, PASCAL
APPLICANT: DALBOW, PASCAL
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: DS/9/389,756
CURRENT FILING DATE: 1999-09-07
EARLIER PETLING DATE: 1999-03-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATCHIN Ver. 2.1
SOFTWARE: PATCHIN VER. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: DALBON, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 44;
         Score 143; DB 3;
Pred. No. 4.8e-14;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143; DB 4;
Pred. No. 4.8e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                   4 PKPQRKTKRNTNRRPQDVKFPGGGOIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Patent No. 6576240
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
                                                                                                                                                                                                                                           Sequence 1, Application US/09389756 Patent No. 6576240 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 1, Application US/08380160 Patent No. 6235284
      92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Hepatitis C virus
  Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: OLIFF & BER
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: VA
COUNTRY: USA
ZIP: 22320
                                                                                                                                                                                                  RESULT 9
US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                          ò
                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08380160

Patent NO. 6235284

GENERAL INFORMATION:
APPLICANT: DAILORN: Michel
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                     ö
                                                                                                          Length 43;
                                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER REDABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FO COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
APPLICATION NUMBER: US/08/380,160
FILING DATE:
                                                                                                        Score 143; DB 4;
Pred. No. 4.7e-14;
1; Mismatches 1;
                                                                                                                                                                                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Beridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET UNBER: WPB 28682
TELECOMMUNICATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Human Hepatitis C Virus
STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BERRIDGE
                                                                                                     Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: OLIFF & BERRI STREET: P.O. Box 19928 CITY: Alexandria STATE: VA COUNTY: USA ZIP: 22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 44 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: (703)836-2787
               TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION: 1..44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: Peptide
LOCATION: 1..44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ANTI-SENSE: NO
STRANDEDNESS:
                                                                                                                                                                                                                                                                                                 RESULT 8
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-380-160-2
                                                                                                                                                                                           õ
                                                                                                                                                                                                                                  QQ
```

ö

Gaps

COMPUTER READABLE FORM:

```
Sequence 23, Application US/08083947

Sequence 23, Application US/08083947

GENERAL INFORMATION:

APPLICANT: Wang, Chang Yi

APPLICANT: Hoseln, Barbara

TITLE OF INVENTION: In Diagnosing and Detecting No. 5639594-A, No. 5639594-

NUMBER OF SEQUENCES:

ADDRESSEE: M. Lisa Wilson

STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                       Score 143; DB 1; Length 61;
Pred. No. 6.8e-14;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 1; Length 61;
Pred. No. 6.8e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATE: US/08/054
FILING DATE: US-SEP-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
      APPLICATION NUMBER: US/07/946,054
FILING DATE: 15-SEP-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    34045
20002
                       FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
ATCONEY, AAGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/POCKET NUMBER: 2000
TELECOMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-nocompage SOFTWARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                            92.98;
                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: LENGTH: 61 amino acids TYPE: amino acid TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                    ; MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE: peptide US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:

APPLICANT: Wang, Chang Yi

TILLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-A, TITLE OF INVENTION: No. 5582968-B Hepatitis

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS: ADDRESS: ADDRESS: ADDRESSE: United Biomedical Inc.

STREET: 25 Davids Dr.

CITY: Hauppauge

STATE: New York

COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 92.9%; Score 143; DB 3; Length 45; Best Local Similarity 92.9%; Pred. No. 4.9e-14; Matches 26; Conservative 1; Mismatches 1; Indels
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 05-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                      NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION: INFORMATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ONGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 9, Application US/07946054
Patent No. 5582968
                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
CUS-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: Peptide LOCATION: 1..45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :ULT 11
```

ô g

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
        ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 26, Application US/08262037
Petent No. 547239
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
CORRESPONDENCE ADDRESS:
ADDRESSE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
          Gaps
        ö
                                                                                                                                                                                                       GENERAL INCORPATION:

APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang YI.

ATILE OF INVENTION: Depetides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                           ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPPRATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
      1; Mismatches
                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                             4 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.9%; Score 143; 92.9%; Pred. No. 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 530
ATORNEY/AGENT INFORMATION:
NAME: WILSON, W. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                             Sequence 3, Application US/08530550 Patent No. 5736321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE: (516)273-2828
TELEFAX: (516)273-1717
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Juery Match
est Local Similarity
Matches 26; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CITY: NEW YORK
STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                          US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                          ò
                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: United Blomedical Inc.
TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
TITLE OF INVENTION: Non-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.9%; Score 143; DB 1; Length 61; 92.9%; Pred. No. 6.8e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSE: UNITED BIOMEDICAL INC.
STREET: 25 Davids Drive
CITY: Hauppauge
STATE: New York
COUNTRY: USA
ZIP: 11788
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN DATA:
APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/67,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990
ATORNEY/AGENT INFORMATION:
NAME: MATAL C. H. LIN
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-40/
TELEPHONE: 212-758-4800
TELEPHONE: 212-758-4800
TELEPHONE: 212-758-4809
TELEPHONE: 212-758-4800
TELEPHONE: 212-758-4809
TELEPHONE: 212-758-4800
TELEPHONE: 312-758-4800
TELEPHONE: 312-758-4800
TELEPHONE: 312-758-4800
TELEPHONE: A1792
INPORMATION FOR SEQ ID NO: 26: SEQUENCE CHARACTERISTICS:
                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                          SYSTEM: PC-DOS/MS-DOS WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 9, Application PC/TUS9308638 GENERAL INFORMATION:
IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 92.98
Matches 26; Conservative
COMPUTER: IBM PC
OPERATING SYSTEM:
SOFTWARE: WORDPER
```

ô

```
us-09-491-146a-28.rai
```

```
RECISTRATION NUMBER: 34,045

REFERENCE/DOCKET NUMBER: 9055

TELECOMMUNICATION INFORMATION:
TELEFAX: 516-273-17128

TELEFAX: 516-273-1717

TELEFAX: 516-273-1717

TELEFAX: 516-273-1717

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
TELEFAX: 516-273-1717

TELEFAX: 516-273-1717

TELEFAX: 516-273-1717

TELEFAX: 516-273-1717

TELEFAX: 516-273-1717

TELEFAX: 516-273-1728

TELEFAX: 516-273-1717

TELEFAX: 516-273-1728

SEQUENCE CHARACTERISTICS:
TELEFAX: 516-273-1728

TELEFAX:
```

Search completed: August 7, 2003, 11:23:51 Job time : 10.6364 secs

Sequence Sequence

```
100
                                                                                  913
          2894
2894
3011
3011
3011
3011
3012
 August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                           451899
     GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                           tal number of hits satisfying chosen parameters:
                                                  US-09-491-146A-28
154
1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                  451899 segs, 118759770 residues
                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                       protein search, using sw model
                                                                    BLOSUM62
Gapop.10.0 , Gapext 0.5
                                                                                                    length: 0
length: 2000000000
```

Title: Perfect score:

Sequence:

protein

ĕ

Run on:

Scoring table:

Searched:

Minimum DB seq Maximum DB seq

Database :

sequence 3, Appli Sequence 16, Appli Sequence 18, Appli Sequence 50, Appli Sequence 51, Appli Sequence 50, Appli Sequence 13, Appli Sequence 13, Appli Sequence 14, Appli Sequence 16, Appli Sequence 17, Appli Sequence 27, Appli

## ALIGNMENTS

44 143 92.9 3012 10 US-09-238-076-2 Sequence 2, Appli 45 143 92.9 3012 11 US-09-995-937-2 Sequence 2, Appli	ALIGNMENTS	RESULT 1 US-10-367-677-1	; Sequence 1, Application US/10367677 ; Publication No. US20030118604A1	GENERAL INFORMATION: APPLICANT: JOILVET MICHEL	; APPLICANT: PENIN, FRANCOIS	; APPLICANT: DALBON, PASCAL ; APPLICANT: LADAVIERE, LAURENT	; APPLICANT: LACOUX, XAVIER	; TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC	; TITLE OF INVENTION: TREATING AN HCV INFECTION	; FILE REFERENCE: 103959	CURRENT FILING DATE: 2003-02-19	; PRIOR APPLICATION NUMBER: US/09/389,756	; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442	; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-05	; NUMBER OF SEQ ID NOS: 11 ; SOFTWARE: Patentin Ver. 2.1	; SEQ ID NO 1	; LENGTH: 44	TYPE: PRT	CHARACTURE CATUS  DIBLICATION INDOBMATION.	; AUTHORS: Ogata, N. et al.	; TITLE: Nucleotide Sequence and Mutation Rate of the H Strain		; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.	; VOLUME: 88	; PAGES: 3392-3396		; RELEVANT RESIDUES: Z TO 45	08-10-30/-0//-1	.90 00 .01 00000	Score 143;
VISO7 PURCOMB Deb:*	/PCT_NEW_PUB.pep:* /USO6_NEW_PUB.pep:* /USO6_PUB.pub.pep:*	/PCTUS_PUBCOMB.pep:* /USOB_NEW_PUB.pep:*	/USO8_PUBCOMB.pep:* /USO9A_PUBCOMB.pep:*	a/US09B_PUBCOMB.pep:*	a/USO9_NEW_PUB.pep:*	a/US10A_PUBCOMB.pep:* a/US10B_PUBCOMB.pep:*	a/US10C_PUBCOMB.pep:*	a/US10_NEW_PUB.pep:*	a/US60_PUBCOMB.pep:*	disted by chance to have a	ď	l score distribution.	·		Description			10	US*I Sequence I, Appli 875-8 Sequence R. Appli	7 Sequence	Sequence 14,	Sequence 78,	Sequence 46,	Sequence 60, A	Sequence 152,	Sequence 152,	Sequence 42,	040*44 Sequence 44, Appl	Sequence 42,	anhac *
pase : Published_Applications_AA:*	1: /cggn2_6/ptodata/1/pubpaa/NCT_REW_POB.pep:* 3: /cgn2_6/ptodata/1/pubpaa/USO6_NEW_POB.pep:* 4: /cgn2_6/ptodata/1/pubpaa/USO6_DUBCOMB.pep:* 5: /cgn2_6/ptodata/1/pubpaa/USO7_NEW_PUB.pep:*	6: /cgn2_6/ptodata/1/pubpaa, 7: /cgn2_6/ptodata/1/pubpaa,	8: /cgn2_6/ptodata/1/pubpaa, 9: /cgn2_6/ptodata/1/pubpaa/	<pre>10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:* 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB_pep:*</pre>	• •	13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:* 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*	` `	16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:	` `	Pred. No. is the number of results predicted by chance to have	score greater than or equal to the score of the result being	and is derived by analysis of the total score distribution.	SUMMARIES		Query Score Match Length DB ID	:	92.9 44 15	92.9 74 10	3 143 92.9 91 9 US=U9=3U8=3U8=1 4 143 92.9 97 10 11S=U9=756=875=8	143 92.9 103 10	143 92.9 108 10	143 92.9 113 10	143 92.9 137 10	143 92.9 138 10	143 92.9 166 11	92.9 166 11	143 92.9 169 II	143 92.9	02.0	TT 60T 6.76 C

Result õ

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 97;
                                                                                                                                                                                                                                  Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cuery Match 92.9%; Score 143; DB 10; Best Local Similarity 92.9%; Pred. No. 1.3e-12; Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                              Score 143; DB 9;
Pred. No. 1.3e-12;
                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE: Suite 701-E Columbia Square STREET: 555 13th. Street, N. W.
                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                      1 PKPORKTKRNAHRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: BRNST, BARBARA G.
RESISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1808-157A
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 60/092,
PRIOR FILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (202)783-6040 INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                  ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: unknown
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: Washington
STATE: D. C.
COUNTRY: U. S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20004
                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             쉽
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                                  Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION: GEERT
APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AGENTS
AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                         Gaps
                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Microsoft Word 6.0 / ASCII text output CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
Pred. No. 5.9e-13;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.9%; Score 143; DB 10; 92.9%; Pred. No. 1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                          COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: 11near
MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Application US/09758308
Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: P.O. BOX 4433
CITY: HOUSTON
92.98;
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 92,9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: TEXAS
                                                                                                                                                                                   US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 3
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                  RESULT 2
                                                                                                   q
                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

```
Sequence 46, Application US/09851138
Publication No. US20020183508a1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                                                                                                                                                                                                                                                                                                                                  polypeptides selected from a
the hepatitis C virus and
                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                          Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 143; DB 10;
Pred, No. 1.6e-12;
                                                                      Score 143; DB 10;
Pred. No. 1.5e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Microsoft Word 6.0 / ASCII CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/09/851,138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                               1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FIRE FILLING FILLING THE PERCECTOR 32
                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: SID nucleic acids and TITLE OF INVENTION: pathogenic strain of TITLE OF INVENTION: applications thereof FILE REFERENCE: 84809A - JAZ
; SEQUENCE DESCRIPTION: SEQ ID NO: 14: US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: 08/836,075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: B 00402225.7
PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                       Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 09-May-2001 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGENTS
                                                                      Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 78
LENGTH: 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-46
                                                                                                                                                                                                                                                                                                  US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                                                               à
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HEPATITIS C VIRUS GENOTYPES
PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                           TITLE OF INVENTION: SID nucleic acids and polypeptides selected from TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REFERENCE: BABODA - JAZ. CURRENT APPLICATION NUMBER: US/09/921,397
CURRENT FILING DATE: 2001-08-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-MAY-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 92.9%; .Score 143; DB 10; Best Local Similarity 92.9%; Pred. No. 1.4e-12; Matches 26; Conservative 1; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 OCT 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 14. Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STOUVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                   PRIOR APPLICATION NUMBER: EP 00402225.7 PRIOR FILING DATE: 2000-08-03
                         Sequence 77, Application US/09921397 Patent No. US20020151484A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Hepatitis C virus 9-09-921-397-77
                                                                                                                                                                                                                                                                                                NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: HOUSTON STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
                                                                      GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-851-138-14
    US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

```
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SECURENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 166;
              Indels
                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
              ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 143; DB 11;
Pred. No. 2.4e-12;
1; Mismatches 1
            Mismatches
                                                1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                    1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.9%; Score 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US/09/878,281
                                                                                                                                                                               Sequence 152, Application US/09899046; Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 152, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 152:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: 152:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.99
Matches 26; Conservative
          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     , MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; MOLECULE TYPE:
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE
                                                                                                                                                          US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-878-281-152
          26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
          Matches
                                                                              g
                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: 08/836,075
FILING DATE: 40TANDAMDATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                               Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 138;
                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                         Score 143; DB 10;
Pred. No. 1.9e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 143; DB 1
Pred, No. 2e-12;
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                    NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                               MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HOLECULE TYPE: peptide SEQUENCE DESCRIPTION: SEQ ID NO: 60: US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 60, Application US/09851138; Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                     LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 138 amino acids TYPE: amino acid
                                                          ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                         92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                     Query Match 92.9°
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                   US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-851-138-60
```

```
TITLE OF INVENTION: New sequences of heparitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIOM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                        Gaps
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 169;
                Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 143; DB 11;
Pred. No. 2.4e-12;
1; Mismatches 1;
              Score 143; DB 11;
Pred. No. 2.4e-12;
                                                      1; Mismatches
                                                                                              1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPORKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: New sequences of TITLE OF INVENTION: genotypes for di NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                              Sequence 42, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 44, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                92.9%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 92.9%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 169 amino acids
amino acid
          Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   , MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         , MOLECULE TYPE: protein US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  inear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-878-281-44
                                                                                                ó
                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                               New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                     ö
                                                                                                                                                                                                                                                                                               TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis of UNBBER OF SEQUENCES 270 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PATENTIAN PC-DOS/MS-DOS SOFTWARE: PATENTIAN PAPLICATION DATA: RPPLICATION DATA: PRIOR APPLICATION DATA: PAPLICATION DATA: PAPLICATION NUMBER: 08/362,455
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 11;
Pred. No. 2.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNAHRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S PKPORKTKRNINRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                   US-09-899-046-42; Sequence 42, Application US/09899046; Publication No. US20030008274A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Sequence 44, Application US/09899046
Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92.9%;
92.9%;
            92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 169 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 92.9
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                 g
```

```
Gaps
                                                          ö
Query Match 92.9%; Score 143; DB 11; Length 169; Best Local Similarity 92.9%; Pred. No. 2.4e-12; Matches 26; Conservative 1; Mismatches 1; Indels (
```

q ά

Search completed: August 7, 2003, 12:01:13 Job time : 14.3636 secs

polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein

genome

genome

genome

genome

genome polyprotein hypothetical prote genome polyprotein genome polyprotein

- hepa

polypeptide genome polypr

polyprotein polyprotein

genome

```
Substitute polyprotein - hepatitis C virus (genotype 1, N7) (fragment)
N.Contains: core protein
N.Contains: core protein
N.Contains: core protein
N.Contains: core protein
A.Variety: genotype 1, N7
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C.Accession: $41349
R.Yan Doorn, L.J.: Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: $41341
A.Accession: $41349
A.Molecule type: genomic RNA
A.Residues: 1-115 cVAN>
A.Kesidues: 1-115 cVAN>
A.Cross -references: EMBL: 229452; NID: 9443866; PIDN: CAA82590.1; PID: 9443867
A.Experimental source: genotype 1, N7
C.Superfamily: hepatitis C virus genome polyprotein
C.Keywords: capsid protein; core protein; polyprotein
F.1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                            N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N9
C;Date: 19-May:1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41351
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41351
A;Accession: S41351
                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (genotype 1, N9) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 139; DB 2; Length 11 Pred. No. 4.2e-11; 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                         A:Experimental source: genotype 1, N9
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-115/Product: core protein *status predicted 
                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 139;
                              PN0677
JQ0881
JQ0881
PC2219
GNWVTC
GNWVCJ
S18030
GNWVCH
                                                                                                                               JC5620
JQ1303
          S18031
S18032
                                                                                                                                                     GNWVJ8
S41356
                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.1%;
ilarity 92.9%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.18;
                                                                                                                                                                                                                                                                                                                                                                         A, Molecule type: genomic RNA
A, Residues: 1-115 <VAN>
A, Cross-references: EMBL: Z29454
                                                              3011
3011
3011
3011
3011
3033
108
782
782
782
787
874
874
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 26; Conserv
 Query Match
 S41349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               olyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           yprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein
                                                                        August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                      283308
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                    al number of hits satisfying chosen parameters:
                                                                                                                               156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                 283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                         Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                      protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S41369
S41368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S21471
S12707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S41349
JQ1584
S19875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PC1284
JQ1925
                                                                                                                                                                           Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S41342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S41351
                                                                                                                                                                                                                                          seq length: 0
seq length: 200000000
                                                                                                                      US-09-491-146A-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115
118
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                                                                                  PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing:
                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score
                                                      OM protein
                                                                                                                                                                                                                                            Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                       Database :
                                                                                                                                           Sequence:
                                                                                                                                                                                                 Searched:
                                                                            Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
```

ö

Gaps

ö

Length 115;

DB 2;

Length 115;

g

ò

```
Cincession: PC2060
Rili, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Bicochem. Blophys. Res. Commun. 199, 1474-1481, 1994
A:Title: Identification of the third major poor poor of hepatitis C virus in France
A:Reference number: PC2060; MUID:94197744; PMID:8147893
A:Recession: PC2060
A:Reference number: PC2060; MUID:94197744; PMID:8147893
A:Molecule type: mRNA
A:Molecule type: mRNA
A:Residues: 1-411 < LIJ>
C:Superfamily: hepatitis C virus genome polyprotein
C:Reywords: ATP: capsid protein: envelope protein: glycoprotein; nonstructural pro
C:Reywords: ATP: capsid protein: anvelope protein: glycoprotein; monstructural pro
C:Reywords: ATP: capsid protein: E2/NSI #status predicted <NPE>.
F:192-383/Product: envelope protein E2/NSI #status predicted <NPE>.
F:196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted to the EMBL Data Library, April 1992
A; Description: A sentitive serodiagnosis of hepatitis C virus infection with two sheference number: $21336
A; Reference number: $21336
A; Accession: $21336
A; Molecule type: genomic RNA
A; Residues: 1-88 < SAD>
A; Cross-references: EMBL: X65548; NID: 959492; PIDN: CAA46517.1; PID: 959493
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) C; Species: hepatitis C virus (c) Species: hepatitis C virus (c) Species: hepatitis C virus (c) Species: Deb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 R; Sato, A.
                                                                         N.Contains: envelope protein E1; nonstructural protein E2/NS1
C;Species: hepatitis C virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 2, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Accession: $41353
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 137; DB 2; Length 411;
Pred. No. 2.6e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 136; DB 2; Len
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Residues: 1-108 < VANN

A:Residues: 1-108 < VANN

A:Cross-references: EMBL: 229456

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: capsid protein; core protein; polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                 polyprotein N1 - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 87.8
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Reference number: S41341
A; Accession: S41353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          521336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 4
$19875
genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
(Contains: core protein; envelope protein 1; nonstructural protein 2; NSI/E2 protein C; Species: hepatitis C virus
A; Variety: isolate JK3
C; Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C; Accession: S19875
R; Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
S: Milted to the EMBL Data Library, September 1991
Scription: Sequence analysis of putative structural regions of Hepatitis C Virus iso
                                                                                                                                                                                                                                                                                                       RESULT 3
JQ1584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01584
JQ01881
JQ0188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            predicted < (Asn) (cova
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       coss-references: GB:X84079; NID:g643119; PIDN:CAA58888.1; PID:g643120
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: core protein; envelope protein; glycoprotein; nonstructural protein; polypro
F;1-191/Product: core protein C *status predicted <CPC>
F;192-389/Product: envelope protein E1 *status predicted <EE1>
F;390-640/Product: envelope protein E2 and nonstructural protein NSI *status predicted <F;196,209,234,305,417,430,448,476,540,556,576,623/Binding site: carbohydrate (Asn) (cove
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       putative cor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nonstructural
                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <MAT4>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Title: Cloning and sequencing of the structural region and expression of A; Reference number: JQ1584; MUID:92300349; PMID:1318944
A; Accession: JQ1584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Scitptudi Sequence analysis of purative structure regiments and sequence analysis of products 18029
A;Accession: S19875
A;Molecule type: genomic RNA
A;Rosidues: 1-782 <480x)
A;Cross-references: EMBL:X61592; NID:g59482; PIDN:CAA43789.1; PID:g59483
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein #status predicted <MAT2>
F;192-383/Product: envelope protein #status predicted <MAT3>
F;384-733/Product: nonstructural protein 2 (fragment) #status predicted <MAT3
                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                             .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 782;
Pred. No. 4.2e-11;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 139; DB 2;
Pred. No. 2.2e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88.5%; Score 138; DB 2;
92.9%; Pred. No. 3.6e-10;
tive 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89.18;
92.98;
92.98;
Similarity 92.9
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      'olecule type: genomic RNA sidues: 1-640 <KUM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Simi
Matches 26;
Best Local
Matches 2
```

g

å

υ

ö

Gaps

g

ò

genome

g

င်

```
genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
C; Species: hepatitis C virus
C; Species: hepatitis C virus
A; Variety: genotype 1, N6
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: $41349
R; Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: $41341
A; Accession: $41348
A; Molecule type: genomic RNA
A; Accession: $41348
A; Molecule type: genomic RNA
A; Residues: 1-108 < VANN
A; Cross-references: EMBL: 229451
A; Experimental source: genotype 1, N6
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein #status predicted < WAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Grand polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N5
C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
C;Accession: 841371
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: 541371
A;Accession: 841371
A;Accession: 841371
A;Residues: 1-112 <VANA
A;Residues: 1-112 <VANA
A;Residues: 1-112 <VANA
A;Repatimental source: genotype 5, N5
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome
C;Superfamily: hepatit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein - hepatitis C virus (genotype 1, NI) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, NI
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 541341
S;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 135; DB 2; Length 11.
Pred. No. 1.3e-10;
Tred. Treches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 135; DB 2; L
Pred. No. 1.3e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA A; Residues: 1-112 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                           ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A,Variety: genotype 2, N6
\ate: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          #text_change 17-Nov-2000
                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Species: hepatitis C virus
A; Variety: genotype 2, N4
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov
C; Accession: S41355
an Doorn, L.J.: Kleter, G.E.M.; Brouwer, J.T.
antred to the EMBL Data Library, January 1994
a, Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cossion: S41357 Captering Section 2012 1900 February 1800 Am Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A.; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA. A.; Reference number: S41341
                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                   N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                       Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 108
                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A.Accession: S4135
A.Molecule type: genomic RNA
A.Residues: 1008 CVAN>
A.Cross-references: EMBL:229458
A.Experimental source: genotype 2, N4
C.Superfamily: hepatitis C virus genome polyprotein
C.Stewords: capaid protein; core protein; polyprotein
F.1-109/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86.5%; Score 135; DB 2; I
89.3%; Pred. No. 1.3e-10;
Live 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Experimental source: genotype 2, N6
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <MAT>
   F;1-108/Product: core protein #status predicted <MAT>
                                                                   Score 135; DB 2;
Pred. No. 1.3e-10;
; Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 135; DB 2;
Pred. No. 1.3e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   (genotype 2,
                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKPOROPKRNTPRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.5%;
89.3%;
                                                                   86.5%;
illarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Cross-references: EMBL:229460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
                                   Query Match
Best Local Similarity
Las 25; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N; Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
S41348
```

ò g ò q

ö

Gaps

ö

ö

Gaps

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                N; Contains: core protein
C; Species: hepatitis C virus
'Tilety: genotype 5, N4
.te: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
.ccession: S41370
R; van boorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LIPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A,Cross-references: EMBL:229473; NID:g443908; PIDN:CAA82611.1; PID:g443909
A,Experimental source: genotype 5, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; corre protein; polyprotein
F;1-114/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule.type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL:229472; NID:g443906; PIDN:CAA82610.1; PID:g443907
A; Experimental source: genotype 5, N3
C; Superfamily: heparitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
A;Cross-references: EMBL:229444; NID:g443850; PIDN:CAA82582.1; PID:g443851
                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length .114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 114;
                                                                                                                                                Length 112;
                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A;Description: Analysis of hepatitis C virus genotypes 1 to A;Reference number: S41341
                                                                                                                                           Score 135; DB 2; L
Pred. No. 1.3e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86.5%; Score 135; DB 2; L
89.3%; Pred. No. 1.4e-10;
tive 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 86.5%; Score 135; DB 2; I
Best Local Similarity 89.3%; Pred. No. 1.4e-10;
Matches 25; Conservative 1; Mismatches 2;
                      A; Experimental source: genotypel, N1
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-112/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                              Query Match 86.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA A; Residues: 1-114 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Accession: 541369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

1 PKPOROPKRNTPRRPQDVKFPGGGGIVG 28

g S

Search completed: August 7, 2003, 11:21:48 Job time : 10.5455 secs

```
ö
                                            C.Species: hepatitis C virus
A;Variety: genotype 5, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41368
R;van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                          A)Cross-references: EMBL:229471; NID:9443904; PIDN:CAA82609.1; PID:9443905
A)Experimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö

    hepatitis C virus (genotype 5, N2) (fragment)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Match 86.5%; Score 135; DB 2; I Local Similarity 89.3%; Pred. No. 1.4e-10; les 25; Conservative 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA A; Residues: 1-114 <VAN>
                             N;Contains: core protein
                                                                                                                                                                                                                                                                                    A; Accession: S41368
  genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
```

## GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-27 156 I PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues

Searched:

127863 al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	ption	hepatitis c						h genome	h genome	h genome	h genome	ı	h genome	h genome	homo sapi	bos t											mycoplasma	triticum ae	saccharomyc	homo sapien	helia	dicty	cas omo	h serine/th
	Description	P27959	001404	001403	P27960	P27961	P26663	P26662	P27958	P26660	P26661	000269	P26664	P29846	P51816	P55106	028653	P06915	P23093	P06472	P03199	007285	096p71	P32653	055112	P04487	049378	P04727	P13185	-	190	$\overline{}$	9493	014738
1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	~	₹*		5	7	<b>v</b> :	ď		S	m	6		72	~	~	-			n		m	~		61			E	£4	z	2	н	z	z
	ΙD	POLG_HCVJ	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVJT	POLG_HCV1	POLG_HCVTW	FMR2_HUMAN	GDF6_BOVIN	2A5D_RABIT	CSP_PLABE	CSP_PLABA	HOR7_HORVU	YLL2_EBV	YLL2_EBVA8	ABA2_HUMAN	MRP_STRSU	FMR2_MOUSE	RNB_HSV11		GDA7_WHEAT	KIN1_YEAST	- 1	11S3_HELAN	UDPG_DICDI	Y848_HUMAN	2A5D_HUMAN
	8	н	-	Н	-	~	Н	7	-		-	-4		-4	-	-	٦	Н	Н	7	-	П	1	~	Н	7		П	-	<del></del>	Н	Н	-	-
	Length	513	520	520	737	73	0	0	0	3033	03	3010	5	3010	1311	436	286	339	347	105	148	148	39	1256	27	161	304	313	1064	3	S)	а	977	602
& Ouery	Match	86.5	9	ė	ů.	ė.	٠	ė.	ė	ė.	è.	'n.	4.	ä	· .	σ.	36.9	4	4	'n.	ش	. 33.3	۳.	۳,	۳.	ώ.	ش	ω.	<u>ښ</u>	2	ς.	ä	32.7	32.4
	Score	135	3	3	Š	3	$\sim$	m	3	3	S	'n	m	3	65	9		54	S			52			Ŋ	51.5	;	ä	ä	51	21	21	51	50.5
Result	No.	-	7	m	4	S	9	7	80	6	10	11	12	13	14	15	16	17	18	19	50	21	22	23	24	25	56	27	28	59	30	31	32	33

P48608 drosophila	P75136 mycoplasma	Q9qyx7 mus musculu	Pl0164 rattus norv	P50477 canavalia e	P10562 canavalia q	P14196 dictyosteli	035144 mus musculu	P10723 brugia mala	Q92777 homo sapien	O15297 homo sapien	P21188 xenopus lae.	
DIA_DROME	YG55_MYCPN	PCLO_MOUSE	PRP2_RAT	CANA_CANEN	CANA_CANGL	AAC2_DICDI	TRF2_MOUSE	SYN_BRUMA	SYN2_HUMAN	P2CD_HUMAN	ATND_XENLA	
-	-	~	ч	-	~	_	٦	Н	Н	-	Н	
1001	204	5038	172	445	445	448	495	548	582	605	277	
32.4	31.7	31.7	31.4	31.4	31,4	31.4	31.4	31.4	31.4	31.4	31.1	
50.5	49.5	49.5	49	49	49	49	49	49	49	49	48.5	
34	35	36	37	38	39	40	4.1	42	43	44	45	

## ALIGNMENTS

Gaps

ö

Indels

Length 520;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annoctation update)
16-OCT-2001 (Rel. 40, Last annoctation update)
6Genome polyprotein [Contains: Capsid protein C (Core protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HCV-KF) (HCV).
Hepatitis C virus (isolate HCV-KF) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLING-93019030; pubMed=1383400;
Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus
"Genomic characterization and mutation rate of hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992)
I. Gen. Virol. 73:2725-2729(1992)
I. TOPMOTENIN BYPELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                             56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                    Score 135; DB 1; Lu
Pred. No. 1.4e-10;
1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Transmembrane; Nonstructural protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; JO1925; JO1925.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
Ffam; PP01543; HCV_capsid; 1.
Pfam; PP01542; HCV_core; 1.
Pfam; PP01539; HCV_env; 1.
Pfam; PP01560; HCV_NN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10687; BAA01529.1; -.
                                                                                                                                                                                                                                                                                                                                                    Query Match 86.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                             520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVHK
Q01403;
                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                          TRANSMEM
CARBOHYD
                                                                                                                                                         CARBOHYD
                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
POLG_HCVHK
                                                                          CHAIN
         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUL-1993 (Rel. 26, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA stage; Flaviviridae;
      NONSTRUCTURAL PROTEIN NS1 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                     (POTENTIAL)
                                                 (POTENTIAL)
                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                             86.5%; Score 135; DB 1; Length 513; 89.3%; Pred. No. 1.3e-10; Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Gycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                     943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLULAR AMINOPEPTIDASE
                                          N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     520 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus (isolate HCV-476) (HCV).
Viruses; SSRNA positive-strand viruses, no
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522, HCV capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                                     55704 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Rel. 26, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D10688; BAA01530.1; -.
                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                     513 AA;
                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q01404;
01-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                          25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVH4
                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                   CARBOHYD
                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                               Query Match
                           TRANSMEM
                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVH4
ID POLG_HCVH4
DT 01-JUL.
DT 01-JUL.
DE GENOME
DE ENVELOI
OC VITUREI
OC VITUREI
OC NEPACII
OC HEPACII
OC HEPACI
                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 2
```

K

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           way
                                          MATRIX PROPEIN (CITATAIN),
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTUBAL PROTEIN NSI/E2 (POTENTIAL).
N'LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      There are no restrictions on 19 as its content is in no
                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN A ND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus (isolate HC-J5) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  having poor
four distinct
REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                       Length 520;
                                                                                                                                                                                                                                                                                                       5; DB 1; Leny-
                                 CAPSID PROTEIN C (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Fill-length sequence of a hepatitis C virus genome having homology to reported isolates: comparative study of four genotypes."
                                                                                                                                                                                                                                                                                     1D2BD0A6FF27349B CRC64;
              CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            737 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     European Bioinformatics Institute. The by non-profit institutions as long
                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                           Score 135;
Pred, No. 1
                                                                                                                                                                                                                                                                                                                                        Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002522; HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NSI.
Pfam; PF01543; HCV_caps1d; 1.
                                                                                                                                                                                                                                                                                       56476 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                     86.5%;
                                                                                                                                                                                                                                                                                                     Juery Match
lest Local Similarity 89.33,
lest Local Similarity 10.33,
lest Local 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                   233
234
305
418
                                                                                                                                                                                                                                                                                     520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJ5
                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                     SEQUENCE
INIT_MET
                                                                                                  TRANSMEM
                                                                                                                    CARBOHYD
                                                                                                                                     CARBOHYD
                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                       윱
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                               CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ./E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capaid protein C (Core protein); Matrix
protein (Envelope protein M); Major envelope protein E; Nonstructural
proteins NS1 and NS2] (Fragment).
Hepatitis C virus (1solate HC-J7) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ő
                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 737;
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Transmenorane; Nonstructural protein.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                       CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;;
                                                                                                                                                                                                                                                                                                                                                                                                   3AF699D82AD501B1
                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                Score 135; DB 1;
NO. 2e-10;
                                                                                                                                                                              POTENTIAL.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. 2e-1
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                   N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                    ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 89.3%; es 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                            86.58;
                                                                                                                                                                                                                                                                                                                                                                                        737
81207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                    ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJ7
                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                               TRANSMEM
                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P27961;
                                                                                                                  CHAIN
                                                                                                                                           CHAIN
                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDB;
PDB;
PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB;
    ö
between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
66n0me polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS5B (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                               MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  86.5%; Score 135; DB 1; Length 737; 89.3%; Pred. No. 2e-10; ive 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                        CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                   Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                      LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (
                                                                                                                                                                                                                                                Nonstructural protein.
                                                                              EMBL; D10077; BAA00971.1; --
InterPro; IPR002522; HCV_capsId.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
Pfam; PP01543; HCV_capsId.
Pfam; PP01547; HCV_core; 1.
                                                                                                                                                                                                       Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                  Polyprotein; Glycoprotein;
                                                                                                                                                                                        Pfam; PF01539; HCV_env; 1
Pfam; PF01560; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 89,3
Matches 25, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             649
737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11105;
                                                                                                                                                                                                                                               Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'n
                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                          INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
POLG_HCVBK
```

g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in oway modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mountied and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      **RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

**MEDLINE-98278466; PubMed-9568891;

**An Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,

**Steinkuehler C., Tomel L., de Francesco R., kuo L.C., Chen Z.;

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Virus: a 2.2-A resolution structure in a hexagonal crystal form.";

"Complex of NS3 protease and NS4A peptide of BK strain hepatitis C

"Virus: a 2.2-A resolution structure in a hexagonal crystal form.";

"L. Protein Sci. 7:837-847(1998).

"I. Protein Sci. 7:837-847(1998).

"I. PROMPTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

"YNDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-ERLATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

"C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral postition, Cys or Thr in Pl and Sea or Ala in Pl'.

"C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                     Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

BODELINE-SP01508B; Pubmed-8661916;

LOVE R.A., Parge H. E., Wickersham J.A., Hostomsky Z., Habuka N.,

MOOMAW E.W., Adachi T., Hostomska Z.;

The crystral structure of hepatitis C virus NS3 proteinase reveals

trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                      SEQUENCE OF 1487-1500.
MEDLINE-9623524; PubMed-8647104;
BOTOWSKI P., Helland M., Oeblmann K., Becker B., Korneteky L.;
Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by CAMP-dependent protein kinase.";
Eur. J. Blochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_caps1d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_core.
                                                                                                                                                              from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1CSJ, 08-NOV-99.
1GX5, 09-APR-02.
1GX6, 10-APR-02.
1QUV, 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A38465; GNWVTC.
1A10; 25-MAR-98.
1JXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .nterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002531;
```

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 21, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22..); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P57); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-91192160; PubMed=1849488;
MEDLINE-91192160; PubMed=1849488;
Maton., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Ohkoshi S., Shimotohno K.;
"Molecular structure of the Japanese hepatitis C viral genome.";
PEBS Lett. 280:325-328(1991).
IFBS Lett. 280:325-328(1991).
IFBS LETT. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
ICATALYTIC COMMONIA WITH SUPPORT OF PROPERTIES OF THE VIRAL RNA REPLICATION.
PRECURSOR POLYPROTECH, COMMONIA WITH ASPO OF Glu in the Viral position, Cys or Thr in Pl and Ser or Ala in Pl'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         neparitis ( VIIIUS (1solate Japanese) (HCV).
VITUSES; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                position, Cys or Thr in Pl and Ser or Als in ri. CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate
                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                         Score 135; DB 1; Length 3010;
Pred. No. 8.9e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-91088550; Pubmed-2175903;
MEDLINE-91088550; Pubmed-2175903;
Maton N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi
Sugimura T., Shimotohno K.;
"Molecular I. Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome
Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                               1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
1112
11120
11120
11133
11134
11161
11161
11171
11198
11198
11198
                                                                                                                                                                                                                                                3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11116;
                                                                                                                                                               1175
1187
1189
1198
                           1122
1129
11139
11149
11162
11165
                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJA
                                                                                                                                                                                                                                  STRAND
SEQUENCE
                                                                STRAND
STRAND
HELIX
TURN
                                                                                                                       TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                           STRAND
g
                                                                                                                                                                                                                                                                                                                                                                                                                               CELLULAR AMINOPERTIDASE.
CASEID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
          Interpro; IPR001490; HCV_NS4b.
Interpro; IPR001490; HCV_NS4b.
Interpro; IPR001086; HCV_NS5a.
Interpro; IPR001096; RNA_POL_DS_PS.
Interpro; IPR007095; RNA_POL_DS_PS.
Interpro; IPR007095; RNA_POL_DS_PS.
Interpro; IPR007094; RNA_POL_DS_PS.
Interpro; IPR007094; RNA_POL_DS_PS.
Interpro; IPR007094; RNA_POL_DS_PS.
Interpro; IPR007094; HCV_CORE; I.
Pfam; PF01509; HCV_NS2; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4a; I.
Pfam; PF01001; HCV_NS4a; I.
Pfam; PF01001; HCV_NS5a; I.
Pfam; PF01001; HCV_NS5a; I.
Pfam; PF01001; HCV_NS5a; I.
Pfam; PF00909; VITAL_RORP; I.
ProDom; P1086062; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                 645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1081
1085
1087
                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HELIX
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
```

Gaps

```
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Legentrian Electric E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko M.A., Lin C., Caron P.R.; "Hepatitis C virus NS3 RNA helicase domain with a bound oligoniclectide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -1- FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-1- FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
               (POTENTIAL)
                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Inchauspe G., Zebedee S., Lee D.H.H., Sugitari M., Nasoff M.,
                                                                                                                                                                                                                                                                                                              Score 135; DB 1; Length 3010;
Pred. No. 8.9e-10;
l; Mismatches 2; Indels
                                                                                                                                                                                                                                                     MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; Pubmed-9493270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657
                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                        (GLCNAC.
                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILLE FILL FILLE F
                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
         N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. MEDLINE-92052256; Pubmed-1658800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-AUG-1992 (Rel. 23, Created)
                                                                           645
645
11 2041
7 2077
2240
2788
\; 327017 M
                                                                                                                                                                                                                                                                                                              86.5%;
                                                                                                                                                                                                                                                                                                                                          88.38;
                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACTIVATION OF NS3
                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID=11108;
         556
576
623
623
2041
2077
2240
2240
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [RNA] (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVH
P27958;
      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 8
         SOFFFFFFF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bloinformatics and the EMBL outstation - the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELORE PROTEIN C (POTENTIAL).
MAJOR ENVELORE PROTEIN C (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEALS.FHELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unterPro; IPR001650; HCV_RGRP.
interPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Fam; PF01543; HCV_caps1d; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEAD.
HCV_caps1d.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF02907; HCV_N33; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; hclicase_C: 1.
Pfam; PF00998; viral_RdRP; 1.
Probom; P0186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                           EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ. HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01539; HCV_env;
PF01560; HCV_NS1;
PF01538; HCV_NS2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SM00487; DEXDc;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            729
1006
1615
1862
2013
PROTEIN C AND MRNA
                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS; S29.001; -. MEROPS; U39.001; -. InterPro; IPR001410; D InterPro; IPR002522; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3010
369
11083
11107
11165
1237
2039
2039
4430
4430
5328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002518;
InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1316
196
209
234
250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2014
347
1083
1107
1165
1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 305
417
423
430
448
532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
```

```
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            772CBB29CCD94753 CRC64;
   (GLCNAC
GELCNAC
GELCNAC
(GLCNAC
 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327142 MW;
                                                                                                                                                                                                           6623

6623

6755

6756

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

6757

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1511
1514
1532
1550
1555
1570
1584
1598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              233
233
233
233
233
233
233
233
302
302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
STRAND
STRAND
HELIX
SEQUENCE
   CARBOHYD
CARBOHYD
CARBOHYD
                                                   CARBOHYD
CARBOHYD
                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
TURN
HELIX
TURN
TURN
STRAND
STRAND
STRAND
                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HELIX
STRAND
HELIX
HELIX
TURN
HELIX
TURN
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                               InterPro; 1PR001521; HCV_core.

InterPro; 1PR0015219; HCV_NS2.

InterPro; 1PR001409; HCV_NS3.

InterPro; 1PR001409; HCV_NS4.

InterPro; 1PR001409; HCV_NS4.

InterPro; 1PR001409; HCV_NS4.

InterPro; 1PR00166; Helicase_C.

InterPro; 1PR00166; HCV_RAP.

InterPro; 1PR0016
 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: ELAND E2. THE NUCLEDCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. PTM: THE STRUCTURAL PROTEINS C, ELAND E2 ARE PRODUCED BY PROTECLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILLARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
PROTESE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                    EMBL. M67463; AAA45534.1; --
PIR, A36814; GNWVCH.
PDB; 1HEI; 25-NOV-98.
PDB; 1ALV; 16-FEB-99.
PDB; 1ALN; 17-JUN-98.
MEROPS; U39.001; --
MEROPS; U39.001; --
INTERPOS; U39.01; --
INTERPOS; UPR00251; HCV_CONTINTERPOS; IPR002521; HCV_CONTINTERPOS; IPR002531; HCV_CONTINTERPOS; IPR002531; HCV_CONTINTERPOS; IPR002531; HCV_CONTINTERPOS; IPR002531; HCV_CONTINTERPOS; IPR002531; HCV_NSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191
383
746
809
1026
1026
1711
1972
2420
369
369
1107
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                347
1083
1107
1165
1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
```

```
TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
g
                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mes by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license&lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

MEDLINE=9204440; PubMed=1658196;

Manchida A., Miyakawa Y., Mayumi M.;

Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";

J. Gen. Virol. 72:2697-2704(1991).

I. Gen. Virol. 72:2697-2704(1991).

I. Gen. Virol. 72:2697-2704(1991).

C. FUNCTION: THE SMALL PROPEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

C. CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24). Last annoctation update)
03-Rel. 41, Last annoctation update)
04-Rel. 41, Last annoctation update)
05-Rel. 41, Contains: Capsid protein C (Core protein) (P22);
05-Rel. (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
05-Protease/helicase NS3 (P70) (Hepacivirin)
05-Protease/helicase NS3 (P70) (Hepacivirin)
05-Protease/helicase NS3 (P70) (Rel. 41, 198); Nonstructural protein NS4 (P6); Nonstructural protein NS5 (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)).
05-Protein NS5 (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)).
05-Protein NS5 (P66) (P70) (RNA-directed NS4 (P4C)).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBONIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                      ö
       Score 135; DB 1; Length 3011; Pred. No. 8.9e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                    2; Indels
                                                                                                                                                                    PRT; 3033 AA
                                    1; Mismatches
       86.5%; Score 135;
89.3%; Pred. No. 8
                                                                1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                 HCV_capsid
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D00944; BAA00792.1; -
                   Local Similarity 89.3
nes 25; Conservative
                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR001490;
                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus.
NCBI_TaxID=11113;
                                                                                                                                                                 POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
       Query Match
                                    Matches
                                                                                          a
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Transmembrane; Coat protein; Envelope protein; Helicase; Arb-binding; Transmembrane; Coat protein; Hydrolase; Serine protease.

Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

Transmembrane; Nonstructural protein Capting Envelopes.

Transmembrane; Nonstructural protein Capting Envelopes.

Transmembrane; Nonstructural protein Capting Envelopes.

Transmembrane; Nonstructural protein NS2 (POTENTIAL).

Transmembrane; Nonstructural protein NS2 (POTENTIAL).

Transmembrane; Nonstructural protein NS4 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 329165 MW; F957F5C1A273BE9E CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 135; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (Rel. 23, Created)
(Rel. 23, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
PPO) IPPO/ I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
POLG_HCVJ8
ID POLG_HCVJ8
AC P26661;
DT 01-AUG-1992 (
DT 01-AUG-1992 (
```

```
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
                                                               CHAIN
                                                                                                                                           CHAIN
                                                                                                                                                       CHAIN
                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way one non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
     Genome polyprotein (Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP35); Envelope glycoprotein E2 (GP68) (GP76) (NS1); Protein P7: NonStructural protein NS2 (P21) (EC 3 4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3 4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3 4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (isolate HC-38) (HCV).
                                                                                                                                                                               MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                               the viral
                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBUNIT: THE UNELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                       Vicology 188:331-341(1992).

-1- FUNCTION: THE SWALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viprecursor polyprotein. commonly with Asp or Glu in the P6 position, Cys or Thr in P and Ser or Ala in P1'.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate
                                                                                                                                                                                                                   *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                -1 - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
 28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LPR00145; HCV_NS2.

LOTO1745; HCV_NS3.

LOTO1745; HCV_NS4a.

LOTEPPO; IPR001496; HCV_NS5a.

INTEPPO; IPR00166; HCV_RRP.

INTEPPO; IPR007095; RNA_POI_DS_PS.

INTEPPO; IPR007094; RNA_POI_PSVIT.

Pfam; PF01543; HCV_COPE; 1.

"m; PF01543; HCV_COPE; 1.

"m; PF015560; HCV_COPE; 1.

"pF01560; HCV_NS".

"pF01538; PCT_NS9.

"pF01538; PCT_NS9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01506; HCV_NS5a; 1.
PF00998; Viral_RdRP; 1
n; PD186062; HCV_NS1; 1
: SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01538; HCV_NS2; 1.
PF02907; HCV_NS3; 1.
PF01006; HCV_NS4a; 1.
PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S29.001; -.
                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                           NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    u39.00
                                                                                                                             Hepacivirus.
                                                                                                                                                                                                                                                genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                 ,55555555555555555555555555542444
```

```
CAPSID PROTEIN C (POTENTIAL).
MATHEX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Bridelpoe protein; Helicase; Arp-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MW; 1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-92295714; PubMed-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
                                                         CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9e-10;
                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                   ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (
                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similary,
hes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                           733
1010
1619
1866
2017
3033
3033
3033
1087
1111
11169
11111
11169
11111
11169
11111
11169
11111
11169
11111
11169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31642;
                                                                                                                                                                                                                         1111
1169
1234
                                                                                                                                                                                             347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJT
```

SIMILARITY) SIMILARITY)

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL). (POTENTIAL). (POTENTIAL)

(POTENTIAL)

(POTENTIAL).
(POTENTIAL).

(POTENTIAL) (POTENTIAL)

(POTENTIAL)

(POTENTIAL)

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NG1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P55); Nonstructural protein NS5A (P65); NONSTRUCTURAL NS5A (P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 134; DB 1; Lengtn JU. Pred, No. 1.2e-09; 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94A1C77435D642BB CRC64;
                                                                CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   326573 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      85.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              llarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
hes 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus.
NCBI_TaxID=11104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA)(N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCV1
P26664;
                                                                ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                           NP_BIND
SITE
                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                     FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCV1
             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                   Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.;
"Modecular cloning of hepatitis C virus genome from a single Japanese
carrier: sequence variation within the same individual and among
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                       (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_copsid, InterPro; IPR002522; HCV_copsid, InterPro; IPR002521; HCV_core. InterPro; IPR002531; HCV_core. InterPro; IPR002518; HCV_NS2. InterPro; IPR002518; HCV_NS3. InterPro; IPR001490; HCV_NS4a. InterPro; IPR001490; HCV_NS4a. InterPro; IPR001490; HCV_NS4a. InterPro; IPR001868; HCV_NS5a. InterPro; IPR007096; HCV_ROSP. InterPro; IPR007095; RNA_pol_DS_FS. InterPro; IPR007094; RNA_pol_DS_VS_INTERPRO; IPR007094; RNA_pol_DS_VS_INTERPRO; IPR007094; RNA_pol_PSV1r. Pfam; PF01543; HCV_core; 1. Pfam; PF01543; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D11168; BAA01943.1; -. PIR; A45573; A45573. PDB; 1A10; 25-MAR-98. PDB; 1JXP; 14-JAN-98. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01006; HCV_NS4a; 1
PF01001; HCV_NS4b; 1
PF01506; HCV_NS5a; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01538; HCV_NS2; 1
PF02907; HCV_NS3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           192
384
730
1007
1616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
```

ö

Gaps

;

Length 3010;

Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last annotation update)
Genome polyprotein (Contains annotation update)
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
"The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";
Virology 188:102-113(1992).

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDOPHORIC. SUGGESTING A POSSIBLE BENERANDE RELATED FUNCTION.

-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate * N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RNA)(N):
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                    (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 131; DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327197 MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3e-09;
                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-92230206; PubMed-1314449;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                  645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVTW
P29846;
                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                          CARBOHYD
                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DATE THE PROPERTY OF THE PROPE
                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities a requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI, PEC (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI, PEC (POTENTIAL).
PROTEASE/HELICASE NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
N-LINKED (GLCNAC. ..)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPRO0252; HCV_capsid.
Interpro; IPR00252; HCV_capsid.
Interpro; IPR00251; HCV_capsid.
Interpro; IPR00251; HCV_capsid.
Interpro; IPR00251; HCV_NS1.
Interpro; IPR002518; HCV_NS2.
Interpro; IPR004109; HCV_NS3.
Interpro; IPR00140; HCV_NS4.
Interpro; IPR00140; HCV_NS4.
Interpro; IPR00166; HCV_NS4.
Interpro; IPR00156; HCV_NS5.
Interpro; IPR00156; HCV_NS5.
Interpro; IPR001069; NN_pol_DS_PS.
Interpro; IPR001004; NN_pol_DS_PS.
Interpro; IPR001004; NN_pol_PSvir.
Pfam; PF01543; HCV_capsid; IPPfam; PF01543; HCV_NS1; IPPfam; PF0156; HCV_NS1; IPPfam; PF0156; HCV_NS2; IPPfam; PF01001; HCV_NS3; IPPfam; PF01001; HCV_NS3; IPPfam; PF01001; HCV_NS3; IPPfam; PF01005; HCV_NS3; IPPfam; PF01005; HCV_NS3; IPPfam; PF01005; HCV_NS3; IPPfam; PF00271; helicase_C; IPPfam; PF00271; DEXDC; IPPFAM; IPPFAM; PF00271; DEXDC; IPPFAM; IPPFAM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  115
191
191
191
1006
1006
1006
1006
1008
1008
11065
11065
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; A39166; GNWVC3.
PDB; 1A1V; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

```
MEDINE-98133924; PubMed-9467002;
Chakrabarti L., Bristulf J., Foss G.S., Davies K.E.;
"Expression of the murine homologue of FMR2 in mouse brain and during
                                                                                                                                                                                                                                                                                                                                                                                                                               FWR2_HUMAN STANDARD; PRT; 1311 AA.
P51815; 043786; 060215; P78407; 013521; 014323; 09UNA5;
01-007-1996 (Rel. 34, Created)
16-007-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Fragila X mental retardation 2 protein (Protein FMR-2) (FMR2P) (0x19 Protein) (Fragile X E mental retardation syndrome protein).
                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gu Y . Shen Y ., Gibbs R.A., Nelson D.L.;
Ifentification of FMR2, a novel gene associated with the FRAXE CCG
repert and CpG island. i.
Nat. Genet. 13:109-113(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [4]
SQUENCE FROM N.A. (ISOFORM 1).
MEDLINE-97446139; PubMed-9299237;
Gecz J., Blelby S., Sutherland G.R., Mulley J.C.;
"Gene structure and subcellular localization of FMR2, a member of new family of putative transcription activators.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human),
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Fetal brain;
MEDLINE=96422267; PubMed=8824884;
Chakrabarti L., Knight S.J.L., Flannery A.V., Davies K.E.;
"A candidate gene for mild mental handicap at the FRAXE fragile
                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Fetal brain, and Placenta;
MEDLINE-96241583; PubMed-8673085;
Gecz J., Gedeon A.K., Sutherland G.R., Mulley J.C.;
"Identification of the gene FMR2, associated with FRAXE mental
                                                                         Length 3010;
327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wang L., Thibodeau S.N.;
Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases
                                                                  Score 128; DB 1;
Pred. No. 7.6e-09;
1; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Dugan-Rocha S.P., Glbbs R.A.;
Submitted (FEB-1997) to the EMBL/GenBank/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 520-1311 FROM N.A. (ISOFORM 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 348-421 FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                 2 KPOROPKRNTPRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                        TISSUE-Brain;
MEDLINE-96241584; PubMed-8673086;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hum. Mol. Genet. 5:275-282(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hum. Mol. Genet. 7:441-448(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. (ISOFORM 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. (ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORM 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     retardation.";
Nat. Genet. 13:105-108(1996).
                                                                  82.1%;
88.9%;
                                                                                              Local Similarity 88.9
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    development.
SEQUENCE
                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
FMR2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    site."
                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                         DDT TO DE TO
                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).
MATIX PROTEIN (POTENTIAL).
MAJOR ENVELODE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTENSE/FILLICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ( POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                     InterPro; IPR002522; HCV_capsid.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002513; HCV_core.
InterPro; IPR002513; HCV_core.
InterPro; IPR002513; HCV_NS1.
InterPro; IPR001409; HCV_NS2.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR00166; HCV_NS5a.
InterPro; IPR007094; RNA_Pol_DS_PS.
InterPro; IPR007094; HCV_NS5a.
InterPro; IPR007094; HCV_NS1.
InterPro; IPR007094; HCV_NS1.
InterPro; IPR007094; HCV_NS1.
InterPro; IPR007094; HCV_NS1.
InterPro; IPR007094; HCV_NS2; InterPro; IPR00706; HCV_NS2; InterPro; IPR00706; HCV_NS3; InterPro; IPR00706; HCV_NS3; InterPro; IPR00706; HCV_NS3; InterPro; IPR00706; HCV_NS4b; InterPro; IPR00706; HCV_NS4b; InterPro; IPR00706; HCV_NS5a; InterPro; IPR00706; HCV_NS1; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LINKED
LINKED
LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKED
                                                                             MEROPS; S29.001; --
MEROPS; U39.001; --
InterPro; IPR001410; DI
InterPro; IPR002522; H
InterPro; IPR002521; H
          1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1083
11103
11107
11107
11230
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
11316
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
```

```
C. -- TISSUE SPECIFICITY: BRAIN (MOST ABUNDANT IN HIPPOCAMPUS AND AMYSDALA), PLACENTA AND LUNG.
C. -- DISBASE: Defects in FMR2 are the cause of FRAXE [MIM:309548], an x-linked form of mental retardation. Loss of FMR2 expression is correlated with FRAXE CGG(N) expension. Normal individuals have 6-35 copies of the repeat, whereas cytogenetically positive, developmentally delayed males have more than 200 copies and show methylation of the associated CPG island.
-- SIMILARITY: BELONGS TO THE AF4 FAMILY.
This SWISS-PROW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelsb-sib.ch).
                                                                                                                                                               IsoId=P51816-3; Sequence=VSP_000211, VSP_000212, VSP_000216;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Missing (in isoform 2, isoform 3 and isoform 4).
/FIId-vSP_000211.
Missing (in isoform 2, isoform 3 and isoform 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                isotorm 4).
/FIId-vSp_000212.
Missing (in isoform 2 and isoform 4).
/FIId-vSp_000213.
                                                                                                                  IsoId=P51816-2; Sequence=VSP_000211, VSP_000212, VSP_000213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Triplet repeat expansion; Disease mutation; Alternative splicing.
VARSPLIC 57 60 Missing (in isoform 2 tecform 2 and
Event=Alternative splicing; Named isoforms=4;
Comment=Additional isoforms seem to exist;
                                                                     IsoId-P51816-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GO; GO:0007275; P:development; TAS. GO; GO:0009405; P:pathogenesis; TAS. Pfam; PF05110; AF-4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF139977; AAD45878.1; JOINED. AF139978; AAD45878.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            U48436; AAC82513.1; -. L76569; AAA99416.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF012624; AAB71534.1;
AF012603; AAB71534.1;
AF012604; AAB71534.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB71534.1
AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAA04822.1
AAB69976.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAD45878.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB71534.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB71534.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X95463; CAA64730.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HGNC:3776; FMR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AF012617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AF139978;
                                                                                                                                           Name=3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30954
                                               Name-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genew;
MIM; 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OR SOLVE STATE STA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88888888888888888888888
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Growth/differentiation factor 6 precursor (GDF-6) (Cartilage-derived
morphogenetic protein 2) (CDMP-2) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Articular cartilage;
MEDLINE-95050604; PubMed-7961761;
Chang S., Hongas J.T., Vukicevic S., Luyten F.P.,
Chang S., Rozak C.A., Reddi A.H., Moos M.;
"Cartilage-derived morphogenetic proteins. New members of the
transforming growth factor-beta superfamily predominantly expressed
in long bones during human embryonic development.";
J. Biol. Chem. 269:28227-28234(1994).
---- SUBUNIT: Homodimer: disulfide-linked (By similarity).
---- SIMILARITY: Belongs to the TGF-beta family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom1;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GROWTH/DIFFERENTIATION FACTOR 6. BY SIMILARITY.
                                                                                                                                                                                                                                                                                                  Score 65; DB 1; Length 1311; Pred. No. 0.73;
N -> K (in isoform 4).
/FTIG-VSP_000214.
Missing (in isoform 4).
/FTIG-VSP_000215.
/FTIG-VSP_000216.
D -> A (IN REF. 3).
A -> V (IN REF. 2).
P -> Q (IN REF. 2 AND 6).
MW; 9F151EC06BB636CB CRC64;
                                                                                                                                                                                                                                                                                                                                                       9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  436 AA
                                                                                                                                                                                                                                                                                                                                                       7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR: B55422, B55452.
HSSP: P18075; IBMP.
InterPro: IPR001819; TGFb.
InterPro: IPR001111; TGFb.N.
Pfam; PF00019; TGF-beta, 1.
Pfam; PF00019; TGF-beta, 1.
Pfam; PF00019; TGF-beta, 1.
PRNITS; PR00438; GFC-SKNOT.
PRODOM; PD000357; TGFb; 1.
PROSTIE; PS00250; TGFB; 1.
PROSTIE; PS00250; TGFB, 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            678 PAPRKEPRPNIPLAPEKKKYRGPGKIV 704
                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRQPKRNTPRRPQDVKFPGGGQIV 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                ¥.
                                                                                                                                                                                                                                                                                                  41.7%;
                                                                                                                                                                                                                                                144739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U13661; AAA61416.1; -.
                                                                                                                                                                                                                                                                                                                              Local Similarity 40.7 tes 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      436
                                                       1311
                                                                                                                                                               195
470
548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bovidae; Bovinae; Bos.
NCBI_TaxID-9913;
     466.
                                                                                                             971
                                                                                                                                                                                                                                             1311 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
     466
                                                       467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GDF6_BOVIN
P55106;
                                                                                                                                                                                           CONFLICT
CONFLICT
SEQUENCE
     VARSPLIC
                                                                                                                                                                                                                                                                                                     Query Match
                                                          VARSPLIC
                                                                                                               VARSPLIC
                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDF6_BOVIN
                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FITTER AND DAY BEAUTIFFE FOR THE PARKET OF THE PROPERTY OF THE
        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

DISULFID

364 433 BY SIMILARITY. 368 435 BY SIMILARITY. 400 1NTERCHAIN (BY SIMILARITY). 27 27 N-LINKED (GLCNAC) (POTENTIAL). 89 89 N-LINKED (GLCNAC) (POTENTIAL). 436 AA; 47873 MW; BO688E12EF8AE91D CRC64;	39 78. Score 63. DB 1. Tonath 136.
BY SIMILARITY. BY SIMILARITY. INTERCHAIN (BY N-LINKED (GLONA N-LINKED (GLONA N-LINKED (GLONA MW; BOG88E12EFGAE	Score 62.
433 435 400 27 89 47873	10 78
364 368 400 27 89 436 AA;	
DISULFID DISULFID DISULFID CARBOHYD CARBOHYD SEQUENCE	Ouery Match
F F F F C S	no

7; Indels Query Match
39.7%; Score 62; DB 1;
Best Local Similarity 41.7%; Pred. No. 0.56;
Matches 10; Conservative 7; Mismatches 7 ò

qq

ö

0; Gaps

Search completed: August 7, 2003, 11:20:01 Job time: 6.90909 secs

Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Database

```
TEALINE STAINWELES;

XX MEDLINE-94201770; PubMed-8151307;

RD Tokita H., Shrestha S.M., Okamoto M., Horikita M.,

RA Tokita H., Shrestha S.M., Okamoto M., Horikita M.,

RI Toka H., Shrestha S.M., Miyakawa Y., Mayumi M.;

RI Toka H., Shrestha S.M., Miyakawa Y., Mayumi M.;

RI Toka H., Shrestha S.M., Miyakawa Y., Mayumi M.;

RI Classification into the third major group.";

CC CLIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

CC PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF

CC PROTEIN C AND MINA (BY SIMILARITY).

DR INTERPO; IPRO02522; HCV_capsid.

DR InterPro; IPRO02521; HCV_core.

DR InterPro; IPRO02531; HCV_LON.

INTERPRO; IPRO02531; HCV_NSI.
                                                                                                                                                         0000000
                                                                                 hepatitis
hepatitis
                                                                                                                                                                                   hepatitis
hepatitis
                                                                                                            hepatitis
hepatitis
                                                                                                                                                hepatitis
hepatitis
                                    hepatitis
                                               hepatitis
                                                      hepatitis
                                                               hepatitis
                                                                        hepatitis
                                                                                                   hepatitis
                                                                                                                              nepatitis
                                                                                                                                                                 hepatitis
                                                                                                                                                                          hepatitis
                                                                                                                                                                                                      hepatitis
                                                                                                                                                                                                               hepatitis
                                                                                                                                                                                                                                hepatitis
                                                                                                                                                                                                                                           hepatitis
                                                                                                                                                                                                                                                  hepatitis
                                                                                                                                                                                                                                                           hepatitis
                                                                                                                                        hepatitis
                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                          008479
008479
0088833
008917448
008917440
008917440
008917440
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
00891744
                  08qp62
                                                                                                                                                                                                                                                          068967
                                                                                                                                                                                                                                                                                                                                 081548 PRELIMINARY; PRT; 415 AA.
081548; 01.NOV-1996 (TrEMBLrel. 01, Created)
01.NOV-1996 (TrEMBLrel. 01, Last sequence update)
01.NOV-1996 (TrEMBLrel. 22, Last annotation update)
01.NOV-1096 (TrEMBLrel. 22, Last annotation update)
Hepatitis Core, env, and part of E2/NSI (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M., Tanaka T., Miyakawa Y., Mayumi M.; "Hepartiss B virus subtypes and hepatitis C virus genotypes patients with chronic liver disease in Nepal."; Hepatology 19:805-809(1994).
                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                081266
068966
P88803
                                                                                                                                                                                           Q9DIT6
Q68798
Q68305
Q81807
Q68951
                                            Q68885
Q81486
                                                              08JP06
081542
                                                                                                          070635
Q8JWM0
                                                                                                                            Q8JWK7
Q8JWK1
                                                                                         Q8JWK0
                                                                                                                                                       068796
                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-94186155; PubMed-8138250;
640
3010
3011
3022
                                                                                                                                                                                                             43
109
782
802
3033
88
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus.
NCBI_TaxID=11103;
                         STRAIN-NE125
081548
                                                                                                                                                                                                                                                                                                                                            O81548 hepatitis c O805043 hepatitis c O81235 hepatitis c O81233 hepatitis c O81231 hepatitis c O81234 hepatitis c O81234 hepatitis c O81236 hepatitis c O8861 hepatitis c O68861 hepatitis c O68863 hepatitis c O68865 hepatitis c O88752 hepatitis c O88752 hepatitis c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hepatitis
                                                                       (without alignments)
281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                             August 7, 2003, 11:05:41; Search time 25.6364 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8v7t0
Q8v7s4
                                                                                                                                                                                  830525
                 Compugen Ltd.
        5.1.6
                                                                                                                                                                                tal number of hits satisfying chosen parameters:
                                                                                                 US-09-491-146A-27
156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
       GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                             protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            081548
08BCW3
081230
081235
081231
081231
081234
081236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q68863
Q68865
Q8V7S2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8V7T0
Q8V7S4
                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                sp_unclassified:*
                                                                                                                                                                                                                                                                                                                sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                               sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                         sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                         sp_organelle:*
sp_phage:*
                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                            sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                            sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                             sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          %
Query
Match Length DB
                                                                                                                                                                                                                                                                    SPTREMBL_23:*
                                                                                                                                                                                                                                                                                              sp_fung1:*
                                                                                                                                                                                                                                                                                                       sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_mhc:*
```

100.0 92.3 91.7 91.7 91.7 91.7 91.7 91.7 91.7

Score

Result Š.

Gaps

ö

us-09-491-146a-27.rspt

```
Othor T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
"The nucleotide sequence of the core region of HCV in Pakistan and Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                      PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D29647; BA31022.1; -.
InterPro; IPR002522; HCV_capsid.

Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ohno T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
                                                                                                                                                                                                                                                                                                                       91.7%; Score 143; DB 12; Length 109; 92.9%; Pred. No. 2.4e-12; 1ve 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91.7%; Score 143; DB 12; Length 109; 92.9%; Pred. No. 2.4e-12; Live 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        109
12235 MW; D5D1BBB10A7E3CAB CRC64;
                                                                                                                                                                                                                               109
12293 MW; 524EBD425A7E3CB2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity 92.9%
                                                                                                                                                                                                                                                                                                                                             Local Similarity 92.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                            109 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE 109 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                               109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Ban2-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Ban2-3;
                                                                                                                                                                                               Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein.
                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     081235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081233
                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SORRERETO
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                    Gaps
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_nenv; 1.
Pfam; PF01560; HCV_NS1; 1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
NON_TER 415 415
SEQUENCE 415 AA; 44650 M3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ohnor T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
"The nucleotide sequence of the core region of HCV in Pakistan and
Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEDCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Theamboonlers A., Bedl K., Chinchal T., Sriponthong M., Chantarasamee P., Poovorawan Y.; Molecular characterization of HCV core region of HCV infection in Thai blood donor."; Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                   100.0%; Score 156; DB 12; Length 415; larity 100.0%; Pred. No. 1.5e-13; Conservative 0; Mismatches 0: Indels 0:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 144; DB 12; Length 119;
Pred. No. 1.9e-12;
1; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   119 AA; 13427 MW; 0EBB321B6CB26A0A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        081230 PRELIMINARY; PRT; 109 AA.
081230; CTEMBLrel. 01, Created)
01-NOV-1996 (TEMBLrel. 01, Last sequence update)
01-NOV-1996 (TEMBLrel. 22, Last annotation update)
01-OCT-2002 (TEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPOROPKRNTPRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAR-2003 (TrEMBLrel. 23, 01-MAR-2003 (TrEMBLrel. 23, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.3%;
ilarity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Watches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-061653a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Banl-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  081230
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
20120
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201200
201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OBBCW3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                      SO TE WORK SO ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DDT DDT DDT BRANK 
                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
```

ö

Gaps

.. 0

109 AA.

```
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
nes 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
     NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-Ban2-2;
                                                                                                                                                               Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q81236
Q81236;
Q1-NOV-1996 (
O1-NOV-1996 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
SEQUENCE
                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                           081234
                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
Q81236
ID Q812
AC Q812
DT 01-N
DT 01-N
                                                                                                                                                                                                                                                                                                      RESULT 8
                                                                                                                                                                                                                                                                                                                  081234
     õ
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-Ball-2;
Ohno T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
Ohno T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;
The nucleotide sequence of the core region of HCV in Pakistan and
Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROFIEN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                  Score 143; DB 12; Length 109;
Pred. No. 2.4e-12;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 143; DB 12; Length 109;
Pred. No. 2.4e-12;
1; Mismatches 1; Indels (
                                                                                                                     109 109
109 AA; 12251 MW; D5D1A6060A7E3CAB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  109 AA; 12261 MW; 461E99060A7E3CA9 CRC64;
                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                           Core protein (Genome polyprotein) (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          109 AA.
                                                                                                                                                                                                                                                                        109 AA
                                                                                                                                                                                                       1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                  InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; I. Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                  91.78;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative 1
                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 01, (TrEMBLrel. 22,
                                                                                                                                                             Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109
                                                                                                                                                                                                                                                                                                               01-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
SEQUENCE
                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                     Q81231
Q81231;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081232
                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
08123
ID 08123
AC 0812
DT 01-N
DT 01-N
DT 01-O
DE CORE
OS HEPA
  RATE COCOCA PART SO
                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Soria
                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Onno T., Mizokami M., Chstopher T., Ohba K., Suzuki K., Roger W.;

"The nucleotide sequence of the core region of HCV in Pakistan and Bangladesh: The geographic characterization of HCV in South Asia.";
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: D29651; BAA21026.1; ---
InterPro; IPR002522; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.7%; Score 143; DB 12; Length 109; llarity 92.9%; Pred, No. 2.4e-12; Conservative 1; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                            Length 109;
                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                    461E99060A7E3CA9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DSD1A6060A7E3CAB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                        Score 143; DB 12;
Pred. No. 2.4e-12;
1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109 AA
                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                    109 109
109 AA; 12261 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           109 109
109 AA; 12251 MW;
                                                                                                                                                                                                                                                                                        91.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 01, (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                              26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
```

```
PKPORQTKRNTPRRPONVKFPGGGOIVG 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                        NCBI_TaxID=42791;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCB1_TaxID=42792;
                                                                                                                                                                                                                                                                                                           STRAIN-IND 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER
NON_TER
SEQUENCE 15
                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                       7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                Q68861
Q68861;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    968863
                                                                 RESULT 11
Q68861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  068863
                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Panigrahi A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
Panigrahi A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
"Genctype determination of hepatitis C virus from Northern India:
"I dentification of a new subtype.";
I. J. Med. Virol. 48:191-198 (1996).

C. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).

REMBL: X91300; CAA62674.1; ---
PROTEIN C AND MANA (BY SIMILARITY).

REMBL: X91300; CAA62674.1; ---
REMBL: X91300; CAA62674.1; ---
REMBL: X91300; CAA62674.1; ---
REMBL: X91300; CAA62674.1; ---
PROTEIN PROMESALI HCV_capsid.

REMBL: PFOLIS42; HCV_capsid.

REMBL: PFOLIS42; HCV_COTE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                   Length 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.7%; Score 143; DB 12; Length 150; 92.9%; Pred. No. 3.4e-12; ative 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                   109 109 MW; AGEDG114583E2EEC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 150 150 150 150 MW; DE75E728FF39B2B3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Capsid protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                            91.7%; Score 143; DB 12;
92.9%; Pred. No. 2.4e-12;
11ve 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-IND 674;
MEDLINE-96432288; PubMed-8835354;
                                                                                                                                                                                                                                                                                InterPro, IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 92.99
Watches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus type 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity .92,9.
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-40363;
                                                                                                                            STRAIN-Ban3-1
                                                                                                                                                                                                                                                                                                                 Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus.
                                                                                                                                                                                                                                                                                                                                  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV CORE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
Q68867
```

ò

```
MEDLINE-9642288; PubMed-8835354;
A MEDLINE-9642288; PubMed-8835354;
A Panigrahi A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
T "Genotype determination of hapatitis C virus from Northern India:
T "Identification of a new subtype.";
J. Med. Virol. 48.191-198(1966).
I. SIBUNIT; THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; X91299; CAA62673.1; -...
R InterPro; IPR002522; HCV_capsid.
R InterPro; IPR002522; HCV_capsid.
R InterPro; IPR002521; HCV_capsid.
R InterPro; IPR01542; HCV_capsid; 1.
                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus type 3g.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Panigrah A.K., Roca J., Acharya S.K., Jameel S., Panda S.K.;
"Genotype determination of hepatitis C virus from Northern India :
Identification of a new subtype.";
J. Med. Virol. 48:191-188(1996).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               150 AA; 16503 MW; 2C38B48C6BB9C16A CRC64;
                                                                   01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) capsid protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 143; DB 12;
Pred. No. 3.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Capsid protein (Genome polyprotein) (Fragment)
HCV CORE.
  150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 29
                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
STRAIN-IND 1452;
MEDLINE-96432288; PubMed-8835354;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91.78;
92.98;
                                          01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                Hepatitis C virus type 3b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26; Conservative
PRELIMINARY;
```

ö

Gaps

Ø

```
"Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBMITT THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GINCORROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AB062202; BAB03396.1, -.
InterPro; IPR00252; HCV_coapsid.
   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                          NON_TER 191 191
SEQUENCE 191 AA; 20740 MW; 64151CA83695EF34 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                    Pfam; PF01543; HCV_caps1d; 1. Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                  91.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
26; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                              TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
                                                                                                                                                       lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kato N.
                                                                                                                  Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8V7T0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positīve-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 150;
                                                                                                                                                                                        Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                   150 AA; 16799 MW; BC2932134026E5FD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16728 MW; BC2D761707FCE1B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
CORE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Capsid protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 143; DB 12;
Pred. No. 3.4e-12;
1; Mismatches 1;
                                                                                                                                                                                        DB 12;
                                                                                                                                                                                      91.7%; Score 143; DB 12; 92.9%; Pred. No. 3.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191 AA.
                                                                                                                                                                                                                                                                                                                                                                                               150 AA.
                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                             1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                     2 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRQTKRNTPRRPQNVKFPGGGQIVG 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-96432288; PubMed-8835354;
EMBL, X91306; CAA62680.1; -.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.78;
92.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus type 3q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1996 (TrEMBLrel.
                                                                                                                                 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1
150 1
150 AA;
                                                                                                                                                                                                         Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=42792;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-IND 1751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER
NON_TER
1
SEQUENCE 15
                                                                                              Polyprotein.
NON_TER
NON_TER
1
SEQUENCE 15
                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV CORE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8V7S2
                                                                                                                                                                                                                                                                                                                                                                                             068865
                                                                                                                                                                                                                                                                                                                                                                                                                 068865;
                                                                                                                                                                                                                                                                                                                                                          TRULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
                                                                                                                                                                                                                                                                                                                                                                          3865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08V7S2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FT
   ò
                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                    DDAR REPRESENTATION OF THE PROPERTY OF THE PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

Gaps

ö

Length 191; 1; Indels

Score 143; DB 12; Pred, No. 4.4e-12;

1; Mismatches

BY A

```
Hepatitis C virus guasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Lesions.";
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
SUBMUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; AB062194; BAB81388.1;
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01542; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           191 AA; 20740 MW; 64151CA83695EF34 CRC64;
                    01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 143; DB 12;
Pred. No. 4.4e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PKPORKPKRNTNRRPODVKFPGGGOIVG 32
     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.7%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
PRELIMINARY;
```

Search completed: August 7, 2003, 11:19:00 Job time: 26.6364 secs

us-09-491-146a-27.rspt

```
TELEFAX: 404-B18-5/52
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
TRNGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        internal
COMPUTER READABLE FORM: MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein uvporteICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                 CIT:.
STATE: Gr
COUNTRY: USA
TTD: 30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FRAGMENT TYPE: ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                       RESULT 1
US-08-921-887-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-921-887-27
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 27, Appl
                                                                                    (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 175,
Sequence 175,
Sequence 6, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 198,
Sequence 232,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 24,
Sequence 104
Sequence 26,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 3,
Sequence 10,
                                                                         August 7, 2003, 11:07:41; Search time 10.5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                      328717
                                                                                                                                                                                                                                                                                                                                /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
/cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
/cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
/cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
/cgn2_6/ptodata/1/1aa/PcTUS_COMB.pep:*
/cgn2_6/ptodata/1/1aa/PcTUS_COMB.pep:*
         5.1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-836-075A-104
US-08-921-887-26
US-08-290-665A-175
PCT-US95-10398-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-389-756-1
US-08-380-160-1
US-08-380-160-1
US-08-083-947-23
US-08-262-033-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-836-075A-14
US-08-324-977-8
US-08-384-616-8
                                                                                                                                                                                                                   al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US94-07088-23
                                                                                                                   US-09-491-146A-27
156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-380-160-6
US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-921-887-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -08-904-686A-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-501-195-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-380-160-2
                                                                                                                                                                                                328717 seqs, 42310858 residues
          GenCore version
Copyright (c) 1993 - 2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                    - protein search, using sw model
                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                        Issued_Patents_AA:*
                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4440000000
                                                                                                                                                                                                                                                                                                                                                                 4·0.0
                                                                                                                                Perfect score:
                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
                                                    OM protein
                                                                                                                                                                                                                                                                                                                        Database :
                                                                                                                                            Sequence:
                                                                                                                                                                                                Searched:
                                                                          Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
No.
```

```
Sequence 27, Application US/08921887
Petent No. 6030771
Petent No. 6030771
Petent No. 6030771
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: KIELDS, HOWARD MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: ENDONCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAINDMENCE FOR DEPRESS:
CORRESPONDENCE ADDRESS:
ADDRESSER: JONES & ASKEW, LLP
                                                                                            1, Appli
2, Appli
2, Appli
16, Appl
1, Appli
                                                                                                                                                                                                                                                                                    App
App
App
App
App
App
App
                                                                                                                                                                                                                                                                                                                                                                                                59,
59,
60,
63,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
        Sequence Seq
                                                                                                                                                                                                                                                                                                                    Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match 100.0%; Score 156; DB 3; Best Local Similarity 100.0%; Pred. No. 4.6e-14; Matches 28; Conservative 0; Mismatches 0;
US-08-836-075A-46
US-08-836-075A-60
US-08-851-195-1
US-08-851-135-2
US-07-681-101-16
US-08-290-65A-155
US-08-290-65A-155
US-08-290-65A-157
US-08-290-665A-157
US-08-290-665A-167
US-08-290-665A-161
US-08-290-665A-161
US-08-290-665A-161
US-08-290-665A-161
US-08-290-665A-161
US-08-290-665A-163
US-08-290-665A-164
US-08-290-665A-164
US-08-290-665A-164
US-08-290-665A-164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3709
TELEFAX: 404-818-3799
```

```
Sequence 26, Application US/08921887

Sequence 26, NOWARD A.

TILLE OF INVENTION: MOSALC PROTEIN AND RESTRICTION

TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME

NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:

ADDRESSEE: JONES & ASKEW, LLP

STREET: 191 Peachtree Street, N.W., 37th Floor

CITY: Atlanta

STATE: GA

COUNTRY: USA

STATE: GA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Tabm PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARR: PAPLICATION DATA:

ADDITOR TRANDED: TR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.1%; Score 139; DB 3; Length 74; 92.9%; Pred. No. 2.1e-11; rive 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                          CITY: HOUSTON
STATE: TEXAS
COMPRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDLUM TYEE: FLOPPY disk
COMPUTER: IBW PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICTOSOTH WORD 6.0 / ASCII tex
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORREY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION POR SEQ ID NO: 104:
FEMERHAL 74 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/921,88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION: NAME: WARREN, WILLIAM L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 89.1
Best Local Similarity 92.9
Matches 26; Conservative
                                                   STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: linear
MOLECULE TYPE: peptide
CORRESPONDENCE ADDRESS: ADDRESSE: ARNOLD, W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-836-075A-104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-921-887-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                          APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
MUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 104, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
APPLICANT: STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: ADD THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 139; DB 3; Length 28; Pred. No. 7.6e-12; 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM PC_Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
ATTORNEY AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
TELEFHONE: 404-818-3799
                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                             PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                               Sequence 24, Application US/08921887
Patent No. 6030771
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY: USA
ZIP: 30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
US-08-836-075A-104
                                                                                                                                                                                                            US-08-921-887-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                         ò
                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
Gaps
                                                                                                   Gaps
                                                                                                                                                                                                                                                                    Sequence 175, Application PC/TUS9510398
GENERAL INFORMATION:
APPLICANT: BOKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H. WUCLEOFIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEBRATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIACNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 191;
                                                        Length 191;
                                                                                                 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 136; DB 5; Length 19
Pred. No. 1.4e-10;
1; Mismatches 2; Indels
                                                        87.2%; Score 136; DB 2;
89.3%; Pred. No. 1.4e-10;
tive 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 32
                                                                                                                                                              5 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 32
                                                                                                                                      1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRICR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-411
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: FLOPPY DISK COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   175:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (212) 758-4800 (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: 345 PARK AVENUE
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 191 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INDIVIDUAL ISOLATE: P8 PCT-US95-10398-175
; INDIVIDUAL ISOLATE: P8
US-08-290-665A-175
                                                  Ouery Match
Best Local Similarity 89.3%
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEX: 421792
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                       RESULT 6
PCT-US95-10398-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZIP: 10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                            g
                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                           DB 3; Length 28;
                                                                                                                                                                                                                                                                                                                                                                       Score 138; DB 3;
Pred. No. 1e-11;
1; Mismatches
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CARRACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPORKPNRNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2026-4116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 175, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/ACENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 201
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                        STRANDEDNESS: not relevant
TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 88.5%;

Best Local Similarity 89.3%;

Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                    Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 191 amino acids
                                                                                                                                                                                                                                                                            FRAGMENT TYPE: internal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: NUCTITLE OF INVENTION: AMILITLE OF INVENTION: COMITITLE OF INVENTION: SEQUENCES: 26 CORRESSIONDENCE SOURCES: 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SS: unknown unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             homosapiens
                                                                                                                                                                                                                 MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NEW YORK
: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                           S
S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JS-08-290-665A-175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM:
                                                                                                                                                                                                                                                                                                                  ORGANISM:
US-08-921-887-26
                                                                                                                                                                                                                                                             ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

us-09-491-146a-27.rai

```
GENERAL INFORMATION:

APPLICANT: DALBON, PASCAL

APPLICANT: DALBON, PASCAL

APPLICANT: JOLIVET, Michel

TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE

TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY

TITLE OF INVENTION: FOR DETECTING THE LATTER

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                       STREET: 3000 K Street, N.W.

CITY: Washington
STATE: 0.C.
COUNTRY: U.S.A.
21P: 20007-5109
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPPRATING SYSTEM: PC-DOS/MS-DOS
SOFTAME: PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
CLASSIFICATION NUMBER: US/09-027015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
FILING DATE: 10-FEB-1997
FILING DATE: 10-FEB-1997
FILING DATE: 10-FEB-1997
APPLICATION NUMBER: JP 8-024045
FILING DATE: US-FEB-1996
ATTONENTY/AGENT INFORMATION:
WANDER: MANDER: MANDER: JP 8-024045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 135; DB 4;
Pred. No. 3.9e-11;
1; Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 2, Application US/08380160 Patent No. 6235284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 05.
TELECOMMUNICATION:
TELEPHONE: (202) 672-5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: OLIFF & BERRIDGE STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-
                                     ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .vcvuogY: linear
; MOLECULE TYPE: peptide
US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
                    CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Alexandria STATE: VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                            AND WHICH CAN BE USED ESPECIALLY
                                                Sequence 6, Application US/08380160
Patent No. 6235284
PAPPLICANT: NO. 6235284
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michal
TITLE OF INVENTION: SYMTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 135; DB 3; Length 34;
Pred. No. 3.1e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READBALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIWAKUMA, TOMIKO
APPLICANT: CHIBA, Yukie
APPLICANT: YAGIAN Shintaro
APPLICANT: HASEGAWA, Akira
TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/POCKET WMBER: WPB 28682
TELECOMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Human Hepatitis C Virus US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 36, Application US/09020846
Patent No. 6322963
GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, Kenjiro
                                                                                                                                                                                                                                                STATE: VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                      USA
22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-020-846-36
                           US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
Michel
SYMMETIC POLYPEPTIDES BELONGING TO THE
HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALL)
FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                          ö
                                                                                                                                           Score 135; DB 4; Length 44; Pred. No. 4e-11;
                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                        4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 530
REIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POTTILE OF INVENTION: FOR DEFECTION
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                           86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
      ; VOLUME: 88
; PAGES: 3392-3396
; DATE: 1991
; RELEVANT RESIDUES: 2 TO 45
US-09-389-756-1
                                                                                                                                           Query Match
Best Local Similarity 89.33
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: 1..45
COTHER INFORMATION: COTHER INFORMATION: GOTHER INFORMATION: US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                              RESULT 11
US-08-380-160-1
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1. Application US/09389756

| Sequence 1. Application US/09389756
| Patent No. 657640
| GENERAL INFORMATION:
| APPLICANT: JOLLPET, MICHEL APPLICANT: DALBON, PASCAL APPLICANT: LADAVIERE, LAURENT
| APPLICANT: LADAVIERE, LAURENT
| APPLICANT: LADAVIERE, LAURENT
| TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTION; FILE REFERENCE: 103959
| CURRENT FILING DATE: 1999-09-07
| CURRENT FILING DATE: 1999-09-07
| EARLIER FILING DATE: 1998-03-05
| NUMBER OF SEQ ID NOS: 11
| SEQ ID NO 1
| LENTH: 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PUBLICATION INFORMATION:
AUTHORS: Ogata, N. et al.
TITLE: Nuclectide Sequence and Mutation Rate of the H Strain Patent No. 6576240
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 86.5%; Score 135; DB 3; Length 44; Best Local Similarity 89.3%; Pred. No. 4e-11; Matches 25; Conservative 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: BETINGG, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ANTI-SENSE: NO FRAGMENT TYPE: N-terminal ORIGINAL SOURCE: ORIGINAL SOURCE: ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: Inear MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KET: Peptide
COCATION: 1..44
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 135; DB 1; Lengtn v. Pred. No. 5.7e-11;
                                                                                                                                              COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPOROPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 PKPQRKTKRNTNRRPQDVKFPGGGIVG 31
                                                                                                                                                                                                        CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORNEY/AGENT NFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/08/530,550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 3, Application US/08530550 Patent No. 5736321 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                 NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Wilson, M. Lisa
REGISSTRATION NUMBER: 34,04:
REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                               (516)273-2828
(516)273-1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
          25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARATERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; MOLECULE TYPE: peptide US-08-083-947-23
                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
STREET: 25 bu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hauppauge
                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                      11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11788
                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 23, Application US/08083947
Patent No. 5639594
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Hosein, Barbara
TITLE OF INVENTION: No. 5639594el Linear and Branched Peptides Effective
TITLE OF INVENTION: in Diagnosing and Detecting No. 5639594-B Hepat
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Wang, Chang Yi
APPLICANT: Wang, Chang Yi
TILEOF INVENTION: Wo. 5582968-1 Branched Hybrid and Cluster
TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-A,
TITLE OF INVENTION: No. 5582968-B Hepatitis
NUMBER OF SEQUENCE: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: United Biomedical Inc.
                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
0
                         DB 3; Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 135; DB 1; Length 61;
Pred. No. 5.7e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUBER: US/07/946,054
FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
                   86.5%; Score 135; DE
89.3%; Pred. No. 4.1e
tive 1; Mismatches
                                                                                                   1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: United Biomedical Inc. STREET: 25 Davids Dr.
                                                                                                                                                                                                                        US-07-946-054-9; Sequence 9, Application US/07946054; Patent No. 5582968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
TELEPHANE: 516-273-1717
TELEPHAN: 516-273-1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: 516-273-1717
INFORMATION FOR SEQ ID NO: 9:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       M. Lisa Wilson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 86.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: peptide 07-946-054-9
                   Query Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hauppauge
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                     ò
                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

ò

Gaps

```
Search completed: August 7, 2003, 11:23:51 Job time : 11.6364 secs
                                                à
                                                                                       qq
                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICART: Chang YI Wang and Barbara Hosein
ATTLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ARTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 135; DB 1; Length 61; Pred. No. 5.7e-11;
                                                                                                                                                                                                   Score 135; DB 1; Length 61;
Pred. No. 5.7e-11;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE/DOCKET NUMBER: 1151-4043 US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                          1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                   4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION DATA:
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/67,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990
ATORNEY/AGENT INFORMATION:
NAME: MATIA C. H. Lin
REGISTRATION NUMBER: 29,323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVE. CITY: NEW YORK STATE: NEW YORK COUNTRY: USA 21P: 10154 CONPUTER READABLE FORM: MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 26, Application US/08262037 Patent No. 5747239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: 421792,
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                     86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 61 amino acids
TYPE: Amino acid
STRANDEDNESS:
                                                                                                                                                                                                   Query Match
Best Local Similarity 89.33
Matches 25; Conservative
TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                                                                           TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
.-08-262-037-26
                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                            ò
```

```
Indels
5;
1; Mismatches
           25; Conservative
Matches
```

APPI APPII APPII APPII APPI

Scoring table:

Searched:

Minimum DB : Maximum DB :

Database :

Perfect score:

Sequence:

protein

Run on:

```
HEPATITIS C VIRUS GENOTYPES PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 23, Appl
Sequence 40, Appl
Sequence 9, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 20, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 3
Sequence 1
Sequence 2
                          Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: 12AAS
COUNTRY: 105A
ZIP: 77210-4433
COMPUTER: 105A
COMPUTER: 105Py disk
COMPUTER: 115Py disk
COMPUTER: 11
US-09-878-281-44
US-09-104-965-2
US-09-104-966-2
US-09-104-949-3
US-09-104-949-3
US-09-851-1138-16
US-09-851-1138-18
US-09-899-046-52
US-09-899-046-54
US-09-899-046-144-05-09-878-281-52
US-09-878-281-52
US-09-973-025-50
US-09-973-025-50
US-09-995-860-50
US-09-995-860-50
US-09-995-860-50
US-09-995-860-50
US-09-995-860-50
US-09-995-860-50
US-09-995-860-50
US-09-995-860-50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-742-659-4

US-09-952-572-9

US-09-929-955-1

US-09-747-419-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: 08/836,075
FILING DATE: cUnknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 104, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUTVER, LIEVER
TITLE OF INVENTION: NEW SEQUENCES OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO: 104:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10
       2894
2985
3011
3011
3011
3011
3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: TEXAS
  US-09-851-138-104
       sequence 17, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 77, Appli
Sequence 14, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 152, Appli
Sequence 152, Appli
Sequence 152, Appli
Sequence 152, Appli
Sequence 44, Appli
Sequence 42, Appli
Sequence 42, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence 104, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appli
                                                                                                                                                                          August 7, 2003, 11:20:11; Search time 14.3636 Seconds (without alignments) 231.506 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | Cyn2_6/ptodata//pubpaa/USO7_PUBCOMB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO7_PUBCOMB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO6_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO6_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO6_PUBCOMB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO7_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO8_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO8_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO8_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO8_NEW_PUB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO8_PUBCOMB.ppp:*
| Cyn2_6/ptodata//pubpaa/USO9B_PUBCOMB.ppp:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       451899
                             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-851-138-104
US-09-851-138-10
US-09-851-138-10
US-09-756-875-8
US-09-756-875-8
US-09-921-397-77
US-09-921-397-77
US-09-921-397-78
US-09-851-138-46
US-09-851-138-60
US-09-851-138-60
US-09-871-138-60
US-09-871-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                               156
1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     451899 seqs, 118759770 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                               protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                       US-09-491-146A-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44
74
91
97
103
1133
1137
1166
1169
1169
```

Result Š

3, Appli 1, Appli 20, Appl 2, Appli

us-09-491-146a-27.rapb

셤

```
Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION.
GENERAL INFORMATION.
TITLE OF INVENTION: PROTEINS
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 14114.034942
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                               Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match

86.5%; Score 135; DB 10;
Best Local Similarity 89.3%; Pred. No. 5.7e-10;
Matches 25; Conservative 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 135; DB 9;
Pred. No. 7.1e-10;
                                             ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MBER: 60/092,339
1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 74 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION
                                                                   STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PatentIn version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86.5%;
ilarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Hepatitis C Virus US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION NUMBER:
PRIOR FILING DATE: 1999-C
NUMBER OF SEQ ID NOS: 5
                                                                                 CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE:
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: DOLIVET, MICHEL
APPLICANT: PENIN, FRANCOIS
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LADAVIERE, LAURENT
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT PAPLICATION NUMBER: US/09/389,756
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
LENGTH 44
                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                         Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                   89.1%; Score 139; DB 10;
llarity 92.9%; Pred. No. 1.8e-10;
Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 135; DB 15;
Pred. No. 3.4e-10;
1; Mismatches 2;
 LENGTH: 74 anino acids
TYPE: anino acid
TYPE: anino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 104:
                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                 5 PKPQRQTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                             Sequence 1, Application US/10367677 ublication No. US20030118604A1 ENERAL INFORMATION:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            86.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                   Ouery Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; RELEVANT KE:
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YPE: PRT
```

ö

ö

à g

```
AL INFORMATION:
APPLICANT: MAERTENS, GEERT
STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES AS PROPHYLACTIC, THERAPEUTIC AND DIAGN
                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 108;
                                                              86.5%; Score 135; DB 10; Length 103;
89.3%; Pred. No. 8.1e-10;
Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 135; DB 10;
Pred. No. 8.5e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                    18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: TEXAS
COUNTRY: USA
ZIP: 7710-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                      Sequence 14, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Sequence 78, Application US/09921397
; Patent No. US20020151484A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 108 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                              Query Match
Best Local Similarity 89.39
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 86.5
Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                           RESULT 7
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                              ô
                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 77, Application US/09921397
Patent No. US20020151484Al
Patent No. US20020151484Al
APPLICANT: HYBRIGENICS
TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and TITLE OF INVENTION: paphogenic strain of the hepatitis C virus and TITLE OF INVENTION: paphogenic strain of the hepatitis C virus and TITLE OF INVENTION: paphogenic strain of the hepatitis C virus and TITLE OF INVENTION: applications thereof FILE REFERENCE: B4803A - JAZ CURRENT APPLICATION NUMBER: US/09/921,397
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 97;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 135; DB 10;
Pred. No. 7.6e-10;
                                                                                 Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Mismatches
                                                                                                                                                                                                                                                     ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
RECISTRATION NUMBER: 1808-157A
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE: peptide US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 25; Conserva
                                                                                                                                                                                                                                                                                    CITY: Wac..
STATE: D. C.
COUNTRY: U. S.
"IP: 20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -09-921-397-77
                                                                   US-09-756-875-8
```

us-09-491-146a-27.rapb

```
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER REDABLE FORM:
MEDIUM TYPE: Floppy disk
                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Microsoft Word 6.0 / ASCII text output
                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                       5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 135; DB 10;
Pred. No. 1.1e-09;
1; Mismatches 2;
                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                     5 PRPQRKTKRNTNRRPQDVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-899-046-152
; Sequence 152, Application US/09899046
; Publication No. US20030008274A1
                                                                                                                                                                                                      ; Sequence 60, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                     APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.5%;
                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 89.37
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
                   25;
                                                                                                                                                             RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-851-138-60
                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
                                                                                               g
                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                             ö
          polypeptides selected from a
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWND
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
                                                                                                                                                                                                                                                                                                                                               Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 137;
                                                                                                                                                                                                                                                                                                                                         Score 135; DB 10; Length 1
Pred. No. 8.9e-10;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 135; DB 10;
Pred. No. 1.1e-09;
TITLE OF INVENTION: SID nucleic acids and polifice of INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof FILE REFERENCE: B4609A. SUS/09/921,397; CURRENT APPLICATION NUMBER: US/09/921,397; CURRENT FILING DATE: 2001-08-02. PRIOR FILING DATE: 2000-08-03. NUMBER OF SEQ ID NOS: 156; SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433 CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQKKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 46, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 137 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STUYVER, LIEVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                       Query Match 86.5%;
Best Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION FOR SEQ ID NO: 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.5%;
89.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                          ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-851-138-46
                                                                                                                                                                                                             SEQ ID NO 78
                                                                                                                                                                                                                                                        TYPE: PRT
                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 9
```

```
COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS.
  Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 166;
                                                                                                                                                                                                                                                                                                                                 Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                             Score 135; DB 11;
Pred. No. 1.3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 135; DB 11;
Pred. No. 1.3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 1.36
1, Mismatches
                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/09/878,28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 152, Application US/09878281
Publication No. US20030032005A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 152: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 89.3%;
Matches 25; Conservative 1
                                                                                                                                                                                                                                                                                                                               86.5%;
89.3%;
                                                                                                                                                      FILING DATE:
INFORMATION FOR SEQ ID NO: 15
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                 Query Match 86.5
Best Local Similarity 89.3
Watches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , MOLECULE TYPE: protein US-09-878-281-152
                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270\,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 169;
                                                                                                                                                                                                                                                                                                              Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                           Indels
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM,
COMPUTER: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                ; DB 11; ; 1.3e-09; ches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.5%; Score 135; DB 11;
89.3%; Pred. No. 1.3e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                  Query Match 86.5%; Score 135; DE
Best Local Similarity 89.3%; Pred. No. 1.36
Matches 25; Conservative 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRQPKRNTPRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 44, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 42, Application US/09878281; Publication No. US20030032005A1; GENERAL INFORMATION:
                                                                                                                            FILING DATE:
INFORMATION FOR SED ID NO: 4
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 89.3
Matches 25; Conservative
                                                                                                                                                                                                                          ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
TITLE OF INVENTION: N
TITLE OF INVENTION: G
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             inear
```

us-09-491-146a-27.rapb

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 86.5%; Score 135; DB 11; Length 169; 9st Local Similarity 89.3%; Pred. No. 1.3e-09; 1tches 25; Conservative 1; Mismatches 2; Indels (
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
PRICIA DATE:
PRICIA DATE:
PRICIA APPLICATION DATA:
INFORMATION FOR SEO ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 anino acids
TYPE: amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRQPKRNTPRRP@DVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

Search completed: August 7, 2003, 12:01:13 Job time : 14.3636 secs

```
August 7, 2003, 11:05:37; Search time 38.5455 Seconds (without alignments) 115.301 Million cell updates/sec
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                                                                    154
1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                             OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                  US-09-491-146A-25
                                                                                                                                                                                                                                                                               BLOSUM62
                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                         Sequence:
                                                                                                                    Run on:
```

A\_Geneseq\_19Jun03: Database :

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

DAT: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT /SIDSI/gcgdata/geneseg/genesegp-embl/AA2001.DAT: /SIDSI/gcgdata/geneseg/genesegp-embl/AA2002.DAT: /SIDSI/gcgdata/geneseg/genesegp-embl/AA2003.DAT: /SIDS1/gcgdata/geneseg/genesegp-emb1/AA1996. /SIDS1/gcgdata/geneseg/genesegp-emb1/AA1989 /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995 /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990 'SIDS1/qcqdata/qeneseq/qeneseqp-embl/AA1999 /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980. /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981. /SIDS1/gcgdata/geneseq/geneseqp·embl/AA1986. /SIDS1/gcgdata/geneseq/geneseqp·embl/AA1987. /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992 /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985 /SIDS1/gcgdata/geneseg/genesegp-embl/AA199; /SIDS1/gcgdata/geneseq/genesegp-emb1/AA1982 /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA198: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA198 /SIDS1/gcgdata/geneseg/genesegp-emb. /SIDS1/gcgdata/geneseq,

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Description	NC mosaic protein	Blood transmiscibl	HCV-S1 full-length	NC mosaic protein	HCV type 3 capsid	Hepatitis C virus	Hepatitis C virus	Hepatitis C virus	Hepatitis C virus
ID	AAY06675	AAR53417	AAE20477	AAY06673	AAB71258	AAR92968	AAR92969	AAR92971	AAR96547
ОВ	20	15	23	20	23	17	17	17	17
% Query Match Length DB	28	3010	3010	28	189	191	191	191	319
% Query Match	100.0	100.0	100.0	94.8	94.8	94.8	94.8	94.8	94.8
Score	154	154	154	146	146	146	146	146	146
Result No.	-	7	٣	4	2	9	7	80	5

New mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species – useful for detecting hepatitis infection in an individual

CN14 fragment of H HCV capsid peptide HCV capsid peptide HCV capsid peptide HCV capsid peptide NOn-A, non-B hepat Hepatitis C virus Human hepatitis C Human hepatitis C HEPATITIS C virus HUMAN HEPATITIS C HUMAN HEPATITIS C HCV core-envelope HCV co	Antigen pHCal01. Antigen pHCbl01. HCV fragment 1 / I HCV fragment 2 / I Branched peptide H Non-A, Non-B Hepat Non-A, Non-B Hepat Hepatitis C virus Antigenic epitope HCV core protein N Hepatitis C virus
16 AAW06487 14 AAR30688 14 AAR30688 15 AAR30689 16 AAR30689 16 AAR30689 17 AAR30689 18 AAR34559 19 AAY2499 10 AAY94409 11 AAY94409 11 AAY94409 11 AAR29534 11 AAR29534 11 AAR29536 11 AAR29542 11 AAR29542 11 AAR29542 11 AAR29542 11 AAR29542 11 AAR29542 11 AAR39545 11 AAR30546 11 AAR30546 11 AAR30546 11 AAR30546 11 AAR30548	
08888008444000000000 11111104	71107707088
WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
01111111110088793333333333333333333333333333333333	₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩

1107863

al number of hits satisfying chosen parameters:

Minimum DB seq length: 0 . Maximum DB seq length: 20000000000

1107863 segs, 158726573 residues

Searched:

## ALIGNMENTS

Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination. NC mosaic protein amino acid fragment C. (USSH ) US DEPT HEALTH & HUMAN SERVICES. AAY06675 standard; Protein; 28 AA 98WO-US17385 97US-0921887 17-JUN-1999 (first entry) Khudyakov YE; WPI; 1999-204671/17. Hepatitis C virus. 25-AUG-1997; 21-AUG-1998; WO9910506-A1 04-MAR-1999, Fields HA, AAY06675; AAY06675 

```
ö
                                         homologous antigenic poptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosalc protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosalc protein and the artificial mosalc protein are useful for detecting a hepatitis infection in an individual. The mosalc gene and protein is also useful for deceting a hepatitis infection and intendividual. The infection, especially hepatitis C. The method of synthesizing the artificial gene and the resulting mosalc protein improve the sensitivity, spectrum of immunoassays. This provides improved detection of hepatitis C virus. Sequences ANYOGGS-633 represent amino acid sequence of each monomer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis; NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR;
                                                                                                                                                                                                                                                                       Gaps
                                   of
                                The invention relates to a mosaic protein, comprising a plurality
                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                             Length 28;
                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                             100.0%; Score 154; DB 20; 100.0%; Pred. No. 1.2e-14;
                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                          Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Leu
                                                                                                                                                                                                                                                                                                                                                                                 AAR53417 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Blood transmiscible NANBHV protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C100 antibody; HCV RNA; NS5 region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       <u>F</u>en
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- His, Arg
                                                                                                                                                                                                                                                                   ó;
       Claim 5; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   non-B hepatitis virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label= Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /label= Asn,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   label- Gln,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /label= Phe,
Misc-difference 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label- Thr,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                            comprising the NC mosaic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference 438
                                                                                                                                                                                                                                                       st Local Similarity atches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 291
                                                                                                                                                                                                                    28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                17-JAN-1995
                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                             very Match
                                                                                                                                                                                                                                                                                                                                                                                                          AAR53417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Non-A,
                                                                                                                                                                                                                                                                                                                                                         RESULT 2
AAR53417
ID AAR5
```

a

```
This sequence is encoded by the genome of a blood transmissible non-A, non-B hepatitis (NANBH) virus. The CDNA sequence was isolated using the primers given in AAQG3500-35. The amplified fragments are used in the detection of hepatitis virus. The target DNA was isolated from serum of chronically infected NANBH patients who were C100 antibody-positive and HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were performed on CDNA and the total human NANBH DNA was constructed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Blood-transmissible non-A non-B hepatitis virus DNA - used detection of hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Page 8-20; 22pp; Japanese.
                                                                                                                                                             Asp
                                                                                                                                                                                                                                                                                                                      Asp
                                                                                                                                                                                                                                                                                                                                                              Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Leu
                          /label= Leu,
Misc-difference 1017
                                                                                                                                                           label- Glu,
                                                                                                                                                                                                                                        label- Met,
                                                                                                                                                                                                                                                                                                                  /label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                      'label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'label- Gly,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /label- Tyr,
/label- Arg,
                                                                                                                    /label- Thr,
                                                                                                                                                                                                label- Met,
                                                                                                                                                                                                                                                                               'label- Asn,
                                                                                                                                                                                                                                                                                                                                                          label- Asn,
                                                                                                                                                                                                                                                                                                                                                                                                 'label- Ala,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /label- Glu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'label - Asp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           label - val,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'label- Cys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                label- Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label- Leu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'label- Arg,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /label- Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92JP-0051885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 2844
                                                                                                                                                                                                                 Misc-difference 1205
                                                                                                                                                                                                                                                                                                                                       Misc-difference 1455
                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 2143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference 2486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 2488
                                                                                                                                      Misc-difference 1056
                                                                                                                                                                              Misc-difference 1201
                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 2462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 2862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 2968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 2989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 2990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1994-163130/20.
N-PSDB; AAQ63499.
                                                                                                                                                                                                                                                                                              Misc-difference 126
                  Misc-difference 759
                                                                                               Misc-difference 103
                                                                                                                                                                                                                                                        Misc-difference 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (KAEN/) KAENNO K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JP06105690-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1;
```

Sequence Query Match

SXS

AAE20477;

RESULT 3 AAE20477

ò g

```
artificial gene and the resulting mosaic protein improve the sensitivity, spectrum of immunoreactivity, and antigen specificity of enzyme immunoassays. This provides improved detection of hepatitis C virus. Sequences AAY06673-683 represent amino acid sequence of each monomer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to a mosaic protein, comprising a plurality of homologous antigenic peptides from different genotypes of a species. The antigenic peptides are from nucleocapsid (NC) proteins. A method for synthesising an artificial gene that encodes the mosaic protein is also provided. The method is designated restriction endonuclease assisted ligation (REAL). The mosaic protein and the artificial mosaic protein are useful for detecting a hepatitis infection in an individual. The mosaic gene and protein is also useful for vaccination against infection, especially hepatitis C. The method of synthesizing the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL; restriction endonuclease assisted ligation; vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mosaic protein, comprising a plurality of homologous antigenic tides from different genotypes of a species - useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 146; DB 20;
Pred. No. 1.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       peptides from different genotypes of a species detecting hepatitis infection in an individual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28
                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (USSH ) US DEPT HEALTH & HUMAN SERVICES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NC mosaic protein amino acid fragment A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTIRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                 PRPORKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB71258 standard; protein; 189 AA.
                                                                                                                                                                                                                                                                     Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     comprising the NC mosaic protein.
                                                                                                                                                                                                                                                              AAY06673 standard; Protein; 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 5; Fig 9; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US17385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0921887
                                                                                                                                                                                                                                                                                                                                                                                                 17-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 95.4.
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Khudyakov YE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1999-204671/17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9910506-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04-MAR-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                               AAY06673;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB71258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Fields
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
AAB71258
                                                                                                                                                                                                RESULT 4
                                                                                                                                                                                                                            AAY06673
                                                                                                                                                                                                                                                                                                                               NAME OF THE OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                              δ
                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The invention relates to nucleic acid construct which comprises an expression cassette including a first polynucleotide region including a 5 non-coding region (NCR) sequence of an RNA virus, a second polynucleotide region including a 3 untranslated region (UTR) sequence of the RNA virus, a second polynucleotide region including a 3 untranslated region (UTR) sequence of the NNA virus and at least a C-terminal portion of a coding sequence of the virus and at least a C-terminal portion of a coding sequence of the virus and at least a C-terminal portion of a coding sequence of the virus and a third polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV-S1 full-length polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleic acid construct; expression cassette; non-coding region; NCR; untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1; Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3010;
                                                                                                                           Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                               ; DB 15;
1.4e-12;
hes 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 100.0%; Score 154; DB 23; Best Local Similarity 100.0%; Pred. No. 1.4e-12; Matches 28; Conservative 0; Mismatches 0;
                                                                                                                           / Match 100.0%; Score 154; C
Local Similarity 100.0%; Pred. No. 1.4
les 28; Conservative 0; Mismatches
                                                                                                                                                                                                                                                              1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                    INST MOLECULAR & CELL BIOLOGY. EHRLICH G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 1; Page 70-81; 81pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAE20477 standard; Protein; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hong WJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV-S1 full-length polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-JUL-2001; 2001WO-IL00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24-JUL-2000; 2000US-220248P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tan YH, Lim SP, Lim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2002-280605/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3010 AA;
                                                                 3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAD33038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200208447-A2.
from 23 clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (MOLE-) INST
(EHRL/) EHRLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31-JAN-2002.
```

ö

Gaps

ö

Sequence

Length 28; Indels

18-NOV-2002

29-AUG-2002.

Heinz FX,

```
AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        El; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                DNA and amino acid sequence of HCV envelope 1 and core proteins -used to determine HCV genotype and as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Match 94.8%; Score 146; DB 17; Length 191; Local Similarity 96.4%; Pred. No. 1.1e-12; les 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus isolate S52 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (USSH ) US DEPT HEALTH & HUMAN SERVICES. (USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (USSH ) US DEPT HEALTH & HUMAN SERVICES.
(USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92969 standard; Protein; 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 4; Page 207; 340pp; English,
                                                                                                                                                                                                                                                                                                           Purcell RH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Purcell
                                                                                                                                                        95WO-US10398.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94US-0290665
                                                                                                                                                                                                                                                                                                        Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                 WPI; 1996-139709/14.
N-PSDB; AAT16642.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bukh J, Miller RH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1996-139709/14
                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9605315-A2.
                                                                                                                                                        15-AUG-1995;
                                                                                                                                                                                                 15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-AUG-1995;
                                                                                                                22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92969;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92969
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          comparising a flavivirus mutant that has a deletion of at least 4 conscentive amino acids from the capsid protein, provided that the conscentive amino acids from the capsid protein, provided that the c-terminal hydrophobic region is not affected by the deletion. The vaccine of the invention has virucide, antinflammatory and hepatotropic activity. The attenuated vaccine, and similar nucleic acid vaccines that encode the mutated capsid protein, are useful for protection against a wide range of flavivirus diseases, e.g. pellow fever, Japanese encephalitis, dengue, classical swine fever, bovine viral diarrhoea and chepatitis C. The specified deletion: (i) produces a reliably attenuated virus that does not revert to virulence; (ii) is exactly defined and does not effect immune responses to important proteins; and (iii) can not generate a non-natural virus by recombination. The mutant viruses cellminate the need to produce large amounts of infectious/virulent viruses and can be produced with less expense. The protective response to infever that to killed vaccines.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This sequence represents a fragment of the capsid protein from Hepatitis c virus (HCV) type 1 described in the disclosure of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Attenuated flavivirus live vaccine, useful for protection against e.g. yellow fever, comprises virus with attenuating deletion of amino acids from the capsid protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This invention describes a novel attenuated flavivirus live vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                   Capsid protein; attenuated vaccine; virucide; antinflammatory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 189;
                                                                                                           hepatotropic; yellow fever; Japanese encephalitis; dengue; classical swine fever; bovine viral diarrhoea; hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94.8%; Score 146; DB 23;
96.4%; Pred. No. 1.1e-12;
11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus isolate HK10 core protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV type 3 capsid protein fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 2; 30pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR92968 standard; Protein; 191
                                                                                                                                                                                                                                                                                                    11-FEB-2002; 2002WO-AT00046.
                                                                                                                                                                                                                                                                                                                                            21-FEB-2001; 2001AT-0000272.
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      02-OCT-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-667064/71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mandl C;
                                                                                                                                                                                                                                                                                                                                                                                       (HEIN/) HEINZ F X. (MAND/) MANDL C.
                                                                                                                                                                        Hepatitis c virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                189 AA;
                                                                                                                                                                                                                 WO200266621-A1.
```

Sequence

hepatitis.

AAR92968;

RESULT 6 AAR92968

ò 음

ö

Gaps

õ g

```
/label= Gly
/note= "amino acid in this position is designated X in
/note= the specification, but codon usage shows that the
only possible amino acid at this pos. is Gly*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "amino acid in this position is designated X in
the specification, but codon usage shows that the
only possible amino acid at this pos. is Gly"
                                                                                                                                                                                                                                                                                                                   Hepatitis C virus; subtype; polymerase chain reaction; amplification;
                                              Gaps
                                           ö
                                                                                                                                                                                                                                                                                     Hepatitis C virus types 7c(8a) isolates VN4 amino acids 1-317.
           Length 191;
                                            Indels
           Score 146; DB 17;
Pred. No. 1.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   144..149
/label- val, Ala, Glu, Gly
                                              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G1y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label- Val, Ala, Glu, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label- Cys, Arg, Ser, Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Val
                                                                            1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                             PCR; primer; probe; antibody; infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /label- Met, Thr, Lys,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'label- Val, Ala, Asp,
                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
Misc-difference 144
                                                                                                                                                                                        $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                 label- Met, Leu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           'label- Ser, Arg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label- Asn, Asp
                                           ö
                                                                                                                                                                                        AAR96547 standard; peptide; 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     label- Phe,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label- Met,
           94.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95EP-0870076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95WO-EP04155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /label= Gly
                                                                                                                                                                                                                                                      (first entry)
      Query Match
Best Local Similarity 96.4°
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Stuyver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1996-251460/25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9613590-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-OCT-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Maertens G,
                                                                                                                                                                                                                                                      10-MAR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09-MAY-1996
                                                                                                                                                                                                                        AAR96547;
                                                                                                                                                        RESULT 9
                                                                                                                                                                                                         Q
                                                                              ö
                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                envelope 1; core protein; HCV genotyping; antibody; vaccine;
                                                                                                       AAR92936-R92987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection.
                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid sequence of HCV envelope 1 and core proteins - determine HCV genotype and as vaccines against HCV infection \,
                            envelope 1 and core proteins - as vaccines against HCV infection
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                             Score 146; DB 17; Length 191;
Pred. No. 1.1e-12;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus isolate DK12 core protein
                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (USSH ) US DEPT HEALTH & HUMAN SERVICES.
(USSH ) US SEC DEPT HEALTH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 4; Page 209-210; 340pp; English.
                             amino acid sequence of HCV determine HCV genotype and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR92971 standard; Protein; 191 AA
                                                                            Claim 4; Page 208; 340pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bukh J, Miller RH, Purcell RH;
                                                                                                                                                                                                                                                                                   Query Match 94.8%;
9est Local Similarity 96.4%;
atches 27; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95WO-US10398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94US-0290665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1996-139709/14.
                                                                                                                                                                                                                                                      191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAT16645
N-PSDB; AAT16643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9605315-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-AUG-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-AUG-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA and a
                                                                                                                                                                                                                                                         Sequence
                               DNA and
                                                used
                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR9297
```

Gaps

ö

Indels

Length 36;

Score 145; DB 16; Pred. No. 2.9e-13; 0; Mismatches 1;

Ā

38888

```
Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-B hepatitis; competitive; inhibition assay.
          HCV infection and to raise antibodies against it.
                                                                                                                                                                   1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                          1 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                         AAR30687 standard; peptide; 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                     94.2%;
ilarity 96.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV capsid peptide No. 23.
                                                                                                                                                                                                                                                                                                                                                                                            (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Jolley ME, Leahy DC,
                                                                               Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1993-018073/02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      virus - for immu
non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9222571-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                               11-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   antibodies.
                                                                                                                                                                                                                                                                                                                                                   AAR30687;
                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                       RESULT 11
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                              The sequences AAR86526-R9657B represent novel sequences isolated from hepatitis C virus subtypes different from subtypes la-c, 2a-d, 3a-f, 4a-f, 5a and 6a. They esp. from the novel subtypes 1d-f, 2e-i, 2k, 2l, 3d, 4k-m, 7a-c or types 9. from the novel subtypes 1d-f, 2e-i, 2k, 2l, 3d, 4k-m, 7a-c or types 9. In or 11. The sequences corresp, to the 5' cuntranslated region (UR), the Core/El, NS4 or NS5 regions of the 5' genome. This sequence represents amino acids 1-317 from the HCV types were isolated from patients with chronic HCV from the Benelux countries, France, Cameroon and Vietnam, because of their abenelux countries, France, Cameroon and Vietnam, because of their abenelux countries, France, Cameroon and Vietnam, because of their abenelux countries, prance, Cameroon and Vietnam, because of their camplified, cloned and genotyped. The 5'UR, Core/El and NSSB regions were sequenced either directly or partially and used to classify the new viruses into (subtypes based on comparison with known sequences. The sequences can also be used to synthesise probes and primars for the sequences can also be used to synthesise probes and primars for the sequences can also be used to synthesise probes and primars for the cantill-HCV antibodies, for HCV typing or to prevent HCV infections.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This sequence is encoded by the oligonucleotide, CN14, and represents the peptide fragment CP14. CP14 is a fragment of the core region of hepatitis C virus (HCV). CP14 may be used in the detection of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                   Hepatitis C virus poly:nucleic acid unique to unidentified sub:type - used to develop probes and primers for new sub:types and vaccines to prevent and treat infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Detection of hepatitis C virus - using oligopeptide fragment of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              94.8%; Score 146; DB 17; Length 319; 96.4%; Pred. No. 1.9e-12; 1ve 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CN14; CP14; core region; hepatitis C virus; HCV; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPORKTKRNTYRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CN14 fragment of HCV core region.
                                                                                                                 Claim 25; Fig 3; 150pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW06487 standard; Protein; 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 1; Page 6; 7pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93JP-0156026.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-JAN-1997 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (IMMO ) IMMUNO JAPAN KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1995-047903/07.
N-PSDB; AAT45055.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        319 AA;
N-PSDB; AAT27958.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JP06327482-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-MAY-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29-NOV-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   core region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW06487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   antibody.
```

a à

Todd JA;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                    The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 3B and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contained in the capsid protein of the virus and themselves contain epitope groups. These peptides can be used in immunosasys for HCV antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 94.2%; Score 145; DB 14; Length 38; Best Local Similarity 96.4%; Pred. No. 3e-13; Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 1F; 66pp; English.
```

RESULT 12 AAR30688

```
The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope antibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                   Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               non-B hepatitis virus; NANBHV; hepatitis C virus; HCV; SNV; NS1; NS2; NS3; antigen; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Non-A, non-B hepatitis virus corel region fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 145; DB 14;
Pred. No. 3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SANW ) SANWA KAGAKU KENKYUSHO CO.
(TOFU ) TONEN CORP.
(TOKR-) 2H TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPORKTKRNTNRRPODVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 1F; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR54065 standard; Protein; 38 AA
                                                                                                                                                                                                                                                   Todd JA;
                                                                                                                                                                                                                    (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                 91US-0714471.
91US-0718052.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92JP-0088140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92JP-0088140.
                                                                                                                                  92WO-0S03635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14-FEB-1995 (flrst entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Leahy DC,
                                                                                                                                                                                                                                                                                  WPI; 1993-018073/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12-MAR-1992;
                                                                  W09222571-A1
                                                                                                                                  29-APR-1992;
                                                                                                                                                                 13-JUN-1991;
20-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JP06141870-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-MAY-1994.
                                                                                                   23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  core; ENV;
                                                                                                                                                                                                                                                     Jolley ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR54065;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR54065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The sequences given in AAR30665-89 represent fragments of the Hepatitis C virus (HCV) amino acid sequence. They represent the beginning of the HCV open reading frame to amino acid 38 and encompass the "common" sequence. These peptides are contained in the capsid protein of the virus and themselves contain epitope artibodies in the diagnosis of non-A, non-B hepatitis (NANBH), and in competitive inhibition assay for detecting HCV specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH;
                                                                                                                                                                               Hepatitis C virus; HCV; open reading frame; "common" sequence; capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH; non-A, non-B hepatitis; competitive; inhibition assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic peptide epitope with sequence encoded by hepatitis-C virus - for immunoassay for antigens for diagnosis of non-A, non-B hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94.2%; Score 145; DB 14; Length 38; 96.4%; Pred. No. 3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKTKRNTYRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Fig 1F; 66pp; English.
                                  AAR30688 standard; peptide; 38 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR30689 standard; peptide; 38 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                          (BAXT ) BAXTER DIAGNOSTICS INC
                                                                                                                                                                                                                                                                                                                                                      92WO-US03635
                                                                                                                                                                                                                                                                                                                                                                                       91US-0714471
91US-0718052
                                                                                                   (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                   HCV capsid peptide No. 24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 96.4'
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV capsid peptide No. 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jolley ME, Leahy DC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1993-018073/02
                                                                                                                                                                                                                                                     Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38 AA;
                                                                                                                                                                                                                                                                                                                                                      29-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                       13-JUN-1991;
20-JUN-1991;
                                                                                                 25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                    WO9222571-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
11-MAY-1993
                                                                                                                                                                                                                                                                                                                     23-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR30689;
                                                                  AAR30688;
```

ò g ö

Gaps

Search completed: August 7, 2003, 11:14:06 Job time : 38.6364 secs

us-09-491-146a-25.rag

```
;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence is that of an antigen used in a novel assay for determining anti-HCV antibody titre. A sample is diluted to an abborbance of not more than 2.0 and the peptide antigen is added. The method is useful for monitoring the effect of interferon treatment on anti-HCV antibody titres.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                      DNA coding a Non-A, non-B hepatitis virus antigen - useful for detecting HCV within serum
                                                                                                                                                                                                                                         ;
                                                                                                                       Hepatitis C virus #4 and #6 genes were isolated (AAQ64068-69). Both genes contain the core, ENV, NS1, NS2 and NS3 regions. A core region fragment is given in AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Determination of the antibody titre against hepatitis C virus antigen - using dilute solution and HCV antigenic peptide(s)
                                                                                                                                                                                                      Query Match
94.2%; Score 145; DB 15; Length 38;
Best Local Similarity 96.4%; Pred. No. 3e-13;
Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV; non-A, non-B hepatitis virus; antigen; immunoassay;
interferon treatment; monitoring; antibody titre; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
94.2%; Score 145; DB 16; Length 40;
Best Local Similarity 96.4%; Pred. No. 3.2e-13;
Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus core antigenic peptide.
                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 11; 22pp; Japanese,
                                                                                                                                                                                                                                                                                                                                                           AAR84559 standard; peptide; 40 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 1; Page 10; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94JP-0082160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94JP-0082160
                                                                                                                                                                                                                                                                                                                                                                                                                06-JUN-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (ARIM/) ARIMA T.
(TORA ) TORAY IND INC.
            WPI; 1994-205028/25.
N-PSDB; AAQ64067.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1995-386063/50.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                              38 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JP07260792-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-OCT-1995.
                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                        AAR84559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                              AAR84559
ઠ
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
```

```
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein
genome polyprotein
genome polyprotein
                                                                                                                                August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polypr
genome polypr
polyprotein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome is denome is denome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
genome
genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome genome
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                        al number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                             154
1 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                  283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                          OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                           US-09-491-146A-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pirl:*
pir2:*
pir3:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR_76:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database :
                                                                                                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                       Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Result
```

		MENTS	rain JT) tein M; h nonstruc	19-may-2000 #text_ Ootsuyama, Y.; Cho,	virus genome from ; PMID:1318627	NID:9221612; PIDN:BAA01943.1; PID:922163	bone (NCBIN:10520) olyprotein nucleotide bindir predicted <cpc></cpc>	F:116-191/Product: envelope protein M #status predicted <epm> F:122-389/Product: major envelope protein E #status predicted <mee> F:330-729/Product: monstructural protein NS1 #status predicted <ns1> F:330-1006/Product: nonstructural protein NS2 #status predicted <ns2 #status="" <ns3="" f:1007-1615="" hepacivirin="" predicted="" product:=""> F:1212-1217/Region: nucleotide-binding motif A (P-loop) F:1312-1317/Region: nucleotide-binding motif B</ns2></ns1></mee></epm>	n NS4a #status predicted n NS4b #status predicted n NS5 #status predicted <	ore 154; DB 1; Length ed. No. 2.8e-13; Mismatches 0; Indels	32 32	(genotype 3, N2) (fragment) 26-Jul-1996 #text_change 17 wer, J.T.
JQ1926 JH0711 S19876 S18031 S18032 PN0677	JO0883 JO0881 PC2219 GNWVCJ S18030 GNWVCH S40770 JC5620 JO1303	ALIGNMENT	s C virus (st envelope pro protein NS4b;	1510n 19 M.; Oot	of hepatitis C MUID:92295714	:D01171;	CBI Dackl genome po drolase; #status	protein moing monding	<pre>l protein l protein l protein</pre>	Sc Pr 0;	PKPQRKTKRNTYRRPQDVKFPGGGQIVG 	irus ion Brou
	88888888		a C env prot	awa,	hep UID:	gy :	rus rus tn c	rote lope ral ural rin e-bi	tura tura tura		DOVK	is c vi e_revis
523 782 782 782	874 874 876 3010 3010 3010 3011 3011 3014		yprotein - hepatitis : capsid protein C; 6 S4a; nonstructural p hepatitis C virus	*sequence_revision N.; Nakagawa, M.; C	5573; M	N> :D11168 HCV-JT	cted rr is C vij protein d prote:	elope properties envelope properties enveloped enveloped enveloped electrocker electrocker enveloped envel	nonstructural nonstructural nonstructural	100.0%; larity 100.0%; Conservative	RNTYRRP(           RNTYRRP(	atit irus 2 uenc r, G
	22000000000000000000000000000000000000		Propreserving	5573 ***********************************	clo B A4	NA <tal : GB rce:</tal 	xtra atit lycoj apsie	adjon non	. ;; ;; , ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ;	Similarity 8; Conser	RKTKI       RKTKI	n - hep protein tis C v teis C v 2p = 3, N 59 # seqi
			rotein capsid a; nons	4557 Kato	ular mber 4557 imin	e: D 3010 nces sou	P, 9	uct: uct: duct: duct: oduc gion	onpo	imil ; C	PKP0 	tein re p atit otyp 1135
•	2 4 4 1 1 1 4 4 4 4 4 4 4 4 4 4 4 4 4 4		ESULT 1 45573 enome polyprotein - 1;Contains: capsid F protein NS4a; nonst ;Species: hepatilis	sion: A	Title: Molecular cloning Reference number: A45573 Accession: A45573 Status: preliminary	Molecule type: DNA Residues: 1-3010 <72 / Cross-references: Gl xp. Experimental source	sequences (Froduce)	31/Produ 39/Produ 29/Produ 306/Produ 1615/Produ 1237/Rec	1616-1862/Product: 1863-2013/Product: 2014-3010/Product:	Query Match Best Local S: Matches 28,	1 20	polyprotins: con the feet to t
3 3 3 3 3 3 4 5 5 5 5 5 5 5 5 5 5 5 5 5	ม พ พ พ ഷ ഷ ഷ ഷ գ գ จ ۲ ค ข ข 0 ப		RESULT 1 A45573 genome poly N;Contains: protein NS C;Species:	C; Accest R; Tanak	A; Title A; Refer A; Acces: A; Status	A; Molecu A; Residu A; Cross A; Experi	A; Note: C; Super: C; Keywol F; 2-115,	F;116-191/Product: en F;192-389/Product: ma F;390-729/Product: nc F;730-1006/Product: F;1207-1615/Product: F;1312-1317/Region: n F;1312-1317/Region: n	F; 1616- F; 1863- F; 2014-	Query Best I Match	Qy Dp	RESULT 2 841359 genome polyprotein - h N;Contains: core prote C;Species: hepatitis C A;Variety: genotype 3, C;Date: 19-May-1994 #s C;Accession: S41359 R;van Doorn, L.J.; Kle

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genorype 3, N3
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: $4150
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Reference number: $41341
A; Reference number: $41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A)Cross-references: EMBL:229461; NID:9443884; PIDN:CAA82599.1; PID:9443885
A)Experimental source: genotype 3, N1
C;Superfamily: hepartists C virus genome polyprotein
C;Seywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA
A; Mostaces: 1-124 CVAN>
A; Residues: 1-124 CVAN>
A; Cross-references: EMBL: 229463; NID: 9443888; PIDN: CAA82601.1; PID: 9443889
A; Experimental source: genotype 3, N3
A; Experimental source: genotype 3, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-124/Product: core protein #status predicted <MAT>
                                                                                                                                                                     A;Cross-references: EMBL:229462; NID:9443886; PIDN:CAA82600.1; PID:9443887
A;Experimental source: genotype 3, N2
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; corre protein; polyprotein
F;1-114/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                      by LiPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LiPA
                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 3, N1) (fragment) N:Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus (genotype 3, N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ργ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             #text_change
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 114;
                                                                                                                                                                                                                                                                                                                                                Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A Variety: genotype 3, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change
C;Accession: S41358
R;van Doorn, L.J; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 l
     Library, January 1994
hepatitis C virus genotypes 1 to 5
                                                                                                                                                                                                                                                                                                                                                                                                     1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 146; DB 2;
Pred. No. 1.5e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                          Score 146; DB 2;
Pred. No. 1.5e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.8%;
96.4%;
                                                                                                                                                                                                                                                                                                                                          Query Match 94.8%;
Best Local Similarity 96.4%;
Matches 27; Conservative
submitted to the EMBL Data Lib
A;Description: Analysis of hep
A;Reference number: $41341
A;Accession: $41359
A;Molecule type: genomic RNA
A;Residues: 1-114 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A Accession: $41358
A Molecule type: genomic RNA
A) Residues: 1-114 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Reference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ٤
```

DB 2; Length 124;

Score 146;

94.88;

Query Match

```
France
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nonstructural prof
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         predicte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nonstructural protein;
                                                                                                                                                                                                                                                                                              Schoole purplement Na. Imparities C. VIII.

S. Species: hepatitis C. VIII.

C. Species: New P. S. P.; Trepo, C.

R.Li, J.S.; VIVVIEKI, L.; Tong, S. P.; Trepo, C.

Blochem. Blophys. Res. Commun. 199, 1474-1481, 1994

A.Title: Identification of the third major genotype of hepatitis C. VIII.

A.Reference number: PC2061, MUID:94197744; PMID:8147893

A.Residues: 1-411 - LIJ.

A.Residues: 1-411 - LIJ.

A.Residues: 1-411 - LIJ.

A.Residues: 1-411 - LIJ.

C. Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural prof. Superfamily: hepatitis C. VIII. genome polyprotein

C. Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural prof. Signer and C. Reywords: ATP; capsid protein El #status predicted cNPE.

F. 196, 209, 234, 305, 325/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (fragment)
N;Contains: core protein; envelope protein; NS1 protein
C;Species: hepatitis C virus
C;Sate: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41288
R;Seelig, R.
submitted to the EMBL Data Library, December 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein S29 (core protein region) - hepatitis C virus (fragment) C;Species: hepatitis C virus C;Species: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000 C;Accession: S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <NS1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 411;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A.Reference number: $41288
A.Accession: $41288
A.Accession: $41288
A.Molecule type: genomic RNA
A.Residues: 1-492 <SEE>
A.Cross-references: EMBL:X76918
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: nepatitis C virus genome polyprotein
C.Superfamily: core protein: core protein: envelope protein: pril-191/Product: core protein #status predicted <CRN>
F:1-191/Product: core protein #status predicted <ENV>
F:373-492/Product: NSI protein (fragment) #status predicted
. 1.6e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 146; DB 2;
Pred. No. 6.3e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Match 94.8%; Score 146; DB 2; Local Similarity 96.4%; Pred. No. 5.3e-13; les 27; Conservative 0; Mismatches 1.
                        Pred. No. 1.6e
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32
                                                                                                                                                32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKPORKTKRNTIRRPODVKFPGGGOIVG
                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                     PKPORKTKRNTIRRPODVKFPGGGOIVG
                                                                                                                                                                                                                                                                                virus
                                                                                                                                                                                                                                                                                polyprotein N2 - hepatitis C
                     96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94.8%;
ilarity 96.4%;
Conservative
                        Best Local Similarity 96.4% Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
nes 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loc
Matches
                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                   депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S21336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S41288
                                                                                                     ö
                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    á
```

셤 ò

```
N;Contains: core protein cylins (genotype 2, NS) (Iragment)
C;Species: hepatitis C virus
A;Variety: genotype 2, NS
A;Variety: genotype 2, NS
C;Date: 19-May 1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41356
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A;Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
                                                                                                                                                                                                                                                                                          Wicontains: core protein
Cispecies: hepatitis C virus
A; Variety: genotype 2, N6
C; Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: 541357
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 2, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ); Pred, No. 1.9e-13; 0; Mismatches 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A)Experimental source: genotype 2, N5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <NAT>
              1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A:Experimental source: genotype 2, .N6
C;Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 145; DB 2;
Pred. No. 1.9e-13;
              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNICRRPQDVKFPGGGQIVG 32
                                                                                               5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Molecule type: genomic RNA
A;Residues: 1-108 <VAN>
A;Cross-references: EMBL:229460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL: 229459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 96.4 Matches 27; Conservative
              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number: S41341
A; Accession: S41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: S41356
           27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                    ò
                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
R;Sato, A.
submitted to the EMBL Data Library, April 1992
A;Description: A sentitive serodiagnosis of hepatitis C virus infection with two cloned A; Beference number: $21336
A;Accession: $21336
A;Coss-references: EMBL:X65548; NID:g59492; PIDN:CAA46517.1; PID:g59493
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
N.Contains: core protein
N.Variety: genotype 2, N2
C;Date: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000
C;Accession: S41333
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Accession: S41353
A;Holecule type: genomic RNA
A;Residues: 1-108 <VANN
A;Cross references: EMBL:229456
A;Experimental source: genotype 2, N2
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein **status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
R;Van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Pescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Reference number: S41341
A;Residues: 1-108 <ANAN
A;Residues: 1-108 <ANAN
A;Cross-references: EMBL:229458
A;Experimental source: genotype 2, N4
C;Superfamily: hepatitis C virus genome polyprotein
C;Reywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <ANAT>
                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94.2%; Score 145; DB 2; Length 108; ilarity 96.4%; Pred. No. 1.9e-13; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 108;
                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                        Score 145; DB 2;
Pred, No. 1.6e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 145; DB 2;
Pred. No. 1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11 PKPQRKTKRNTSRRPQDVKFPGGGQIVG 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
nes 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
```

ò ş

ö

```
Search completed: August
Job time: 9.54545 secs
                                                                                                                                                                                                                                                   . a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: S4134 B sequence_revision 26-Jul-1996 #text_change 17-Nov-2000 C;Accession: LJ: Kleter, G.E.M.; Brouwer, J.T. R;Van Doorn, LJ: Kleter, G.E.M.; Brouwer, J.T. A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA. A:Reference number: S4134 A:Reference number: S4134 A:Reference number: S4134 A:Residues: 1-108 </ANN A:Residues: 1-108 </ANN A:Residues: 1-108 </ANN A:Coss.references: EMBL:Z29451 A:Experimental source: genotype 1, N6 C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C virus genome polyprotein C;Reywords: capsid protein; core protein; polyprotein F;1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                %):Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 5, N5
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (genotype 1, NI) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, NI
C;Date: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rivan Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A.Reference number: S41341
A.Recession: S41341
A.Recession: S41341
A.Recion: L.J.2 cVAN
A.Residues: 1-112 cVAN
A.Residues: 1-112 cVAN
A.Residues: C.E.M. NID:9443850; PIDN:CAA82582.1; PID:9443851
A.Experimental source: genomy c virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: core protein; core protein; polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA
A; Reference number: S4131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 145; DB 2; Length 112;
Pred. No. 2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A Accession: S41371
A Molecule type: genomic RNA
A; Readues: 1-112 C4NA>
A; Residues: 1-112 C4NA>
A; Coss - references: EMBL: 229474
A; Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; setatus predicted <AMT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 145; DB 2;
Pred, No. 1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTKRNTYRRPODVKFPGGGOIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94.2%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Accession: 541371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

```
ö
                                                                                          ö
                                                                                                                                                                                                                                                                                                        N;Contains core protein
C;Species: hepatitis C virus
A;Variety: genotyte 5, N4
C;Date: 19-May_1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1-114 <VAN>
A; Cross-references: EMBL: 229473; NID:9443908; PIDN:CAA82611.1; PID:9443909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                             LIPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                          ö
                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ρλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 114;
                                             Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ç
                                                                                                                                                                                                                                                                                                                                                                                                     R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994 A: Description: Analysis of hepatitis C virus genotypes 1 A; Reference number: S41341 A; Accession: S41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A.Experimental source: genotype 5, N4 (5) Superfamily: hepatitis C virus genome polyprotein C.Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted 
F;1-112/Product: core protein #status predicted <MAT>
                                             DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94.2%; Score 145; DB 2; 96.4%; Pred. No. 2e-13; iive 0; Mismatches
                                                                 2e-13;
                                                                   Pred. No. 2e-1
0; Mismatches
                                                                                                                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7, 2003, 11:21:47
                                          Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 96.4
nes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                            C; Accession: 541370
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec

US-09-491-146A-25 154

Perfect score:

1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

Searched:

127863 tal number of hits satisfying chosen parameters:

127863 seqs, 47026705 residues

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_41:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		59 hepatitis c 04 hepatitis c				c	c	၎	q	q	c	h genome	옫		93 bos taurus		77 rattus norv	00 m nuclear r				herpes			-	n9 mus musculu	73 xenopus lae	55 manduca sex			~	30 homo sapien
Desc	00026	P27959	0014	P279	P27	P266	P266	P27	P266	P266	P298	P266	P099	P23188	028193	P107	P233	000600	080361	P50894	P23821	P28276	P1557	P38261	031162	Q8r0n9	P2157	P4815	P0236	049937	P75278	015
ID	POLG_HCVJT	POLG_HCVJ2	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCVTW	POLG_HCV1	FURI_HUMAN	FURI_MOUSE	FURI_BOVIN	SYN_BRUMA	FURI_RAT	NCO3_MOUSE	RK4_TOBAC	RS7_FUGRU	RS7_HUMAN	IE63_HSV2H	B3AT_CHICK	YBV2_YEAST	RL16_SPICI	ZDH1_MOUSE	YB1_XENLA	RS7_MANSE	RS7_XENLA	RK4_SPIOL	YF08_MYCPN	PDPK_HUMAN
DB	٦.		·	7	7	-	H	~	~	-	-	-	-	н	-	~	-	-	<b>-</b>	Н	-	-	~	-	-	-	Н	-	-	-4	٢	~
% Query Match Length DB		513						301				30	σ	793	797	548	793	1398	282	194	194	512	922	753	137	485	303	190	194	293	509	226
% Query Match	0 .	2. 20	4	4	4	4	4	4	4	4	9	9	39.0	ů.	S.		۳.	۳.	•	ij.	÷	ä	ή.	<u>.</u>		。	30.2	φ.	•	œ.	ė.	29.9
Score	154	145	145	4	4	4	4	4	145	4	$\sim$	137	9	52	54	51	51	51	20	49	49	49	4	47.5	47	4	46.5	46	46	46	46	4 6
Result No.		74 F	4	5	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		56	27		29		31	32	33

MEROPS; S29.001; -. MEROPS; U39.001; -. InterPro; IPR001410; DEAD.

Q9z2a0 mus musculu O55173 rattus norv	Q92848 chlamydia p P51816 homo sapien P19711 bovine vira	Q9vrj8 drosophila Q10101 schizosacch P48431 homo sapien	P48432 mus musculu P54231 ovis aries P18447 murine coro	P03416 murine coro
PDPK_MOUSE PDPK_RAT	RNR_CHLPN FMR2_HUMAN POLG_BVDVN	U139_DROME RS7_SCHPO SOX2_HUMAN	SOX2_MOUSE SOX2_SHEEP NCAP_CVM3	ncap_cvma5
	~ ~ ~			7
559	676 1311 3988	108 195 317	319 320 454	454
29.9	29.9 29.9 29.5	29.5 29.5 29.5	29.2 29.2 29.2	29.5
44.	46 45.5	4 4 4 សសស	4 4 4 2 3 3 3	<b>4</b> .
9.00	37 38 38	399. 4140	4 4 4 2 8 4	4.5

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
Makazawa T., Hijkata M., Ishimura Y., Shimotohno K.;
Molecular cloning of hepatitis C virus genome from a single Japanese rarier: sequence variation within the same individual and among T infected individuals.

Virus Res. 23:39-53(1992)

L'EUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

SS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

CHARALTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral position, Cys or Thy in Pl and Ser or Ala in Pl'.

POSSILION: Cys or Thy in Pl and Ser or Ala in Pl'.

C -1-GATALTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                          01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last annotation update)
66-2003 (Rel. 42, Last annotation update)
66-2004 (Rel. 42, Last annotation update)
67-2005 (Rel. 42, Last annotation update)
67-2006 (NSI); Protein P7; Nonstructural protein NS2 (PS2);
67-2006 (NSI); Protease/helicase NS3 (P70) (Hepacivirin NS2 (P21))
68-3-4-21-3; Protease/helicase NS3 (P70) (Hepacivirin NS2 (P21))
68-3-4-21-39; Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
69-2006 (RNA-directed RNA polymerase) (EC 2.7.7.48)].
60-2006 (RNA-directed RNA polymerase) (RC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
       PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D11168; BAA01943.1; -.
       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=31642;
POLG_HCVJT
Q00269;
                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
```

'n

STANDARD;

POLG\_HCVJ2 P27959;

POLG\_HCVJ2

RESULT 2

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN SI/EZ (POTENTIAL).
NON STRUCTURAL PROTEIN NSI/EZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NON STRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
CHARGE RELAX SYSTEM (BY SIMILARITY).
CHARGE RELAX SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 154; DB 1; 100.0%; Pred. No. 1.3e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.

-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.)
N-LINKED (GLCNAC.)
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                            InterPro: IPR002866; HCV_NSSA.
InterPro: IPR002166; HCV_RGRP.
InterPro: IPR0071095; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_PSVIr.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01559; HCV_EN; 1.
Pfam; PF01539; HCV_NS1; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF0100998; Viral_RGRP; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ресн вох
                                                              InterPro; IPR002531; H
InterPro; IPR002518; H
InterPro; IPR004109; H
InterPro; IPR000445; H
InterPro; IPR001490; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
FTTTTTTTTTTTTT
```

```
Leg of B1.

.cled Institutions as .

.statement is not remo.

.rea a license agreement (Se.
email to license agreement (Se.

.nterPro; IPRO0252; HCV_capsid.

DR InterPro; IPRO0252; HCV_capsid.

DR Pfam; PPO1543; HCV_capsid.

DR Pfam; PPO1545; HCV_RS; I.

CHAIN RW POLYPOTECH; COAT PROPERING C POPERTIAL).

CHAIN 192 383 MAJOR BRUELOPE PROPERIN B1 (POTENTIAL).

"WEM 347 569 POTENTIAL).

"WEM 347 569 POTENTIAL).

"WEM 347 569 POTENTIAL).

"O 209 209 N'LINKED (GLCANC. ...) (POTENTIAL).

234 234 234 N'LINKED (GLCANC. ...) (POTENTIAL).

235 234 234 N'LINKED (GLCANC. ...) (POTENTIAL).

236 196 196 N'LINKED (GLCANC. ...) (POTENTIAL).

237 234 234 N'LINKED (GLCANC. ...) (POTENTIAL).

238 234 N'LINKED (GLCANC. ...) (POTENTIAL).

239 234 N'LINKED (GLCANC. ...) (POTENTIAL).

231 233 N'LINKED (GLCANC. ...) (POTENTIAL).

234 234 N'LINKED (GLCANC. ...) (POTENTIAL).

235 N'LINKED (GLCANC. ...) (POTENTIAL).

236 N'LINKED (GLCANC. ...) (POTENTIAL).

237 N'LINKED (GLCANC. ...) (POTENTIAL).

238 234 N'LINKED (GLCANC. ...) (POTENTIAL).

239 234 N'LINKED (GLCANC. ...) (POTENTIAL).

231 233 N'LINKED (GLCANC. ...) (POTENTIAL).

234 N'LINKED (GLCANC. ...) (POTENTIAL).

235 N'LINKED (GLCANC. ...) (POTENTIAL).

236 N'LINKED (GLCANC. ...) (POTENTIAL).

237 N'LINKED (GLCANC. ...) (POTENTIAL).

238 231 N'LINKED (GLCANC. ...) (POTENTIAL).

239 231 N'LINKED (GLCANC. ...) (POTENTIAL).

230 CONDORTED (GLCANC. ...) (POTENTIAL).

231 N'LINKED (GLCANC. ...) (POTENTIAL).

232 N'LINKED (GLCANC. ...) (POTENTIAL).

233 233 N'LINKED (GLCANC. ...) (POTENTIAL).

241 N'LINKED (GLCANC. ...) (POTENTIAL).

252 N'LINKED (GLCANC. ...) (POTENTIAL).

253 233 N'LINKED (GLCANC. ...) (POTENTIAL).

254 COTEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license eigenement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
60-ORD (Rel. 40, Last annotation update)
Envelope glycoprotein [Contains: Capsid protein C (Core protein (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NSI)] (Fragment).
Hepatitis C virus (isolate HC-J2) (HCV).
Hepatitis C virus (isolate HC-J2) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genotypes.",
Virology 188:331-341(1992)
-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
-!- FUNCTION: THE SMALL PROTEINS NS2B, NS2B, NS4B AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- SUBUNIT: THE VIRICN OF THIS VIRUS IS A NICLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.

BEDLINE-92330232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Trill-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
                                                                                                                                                                                                                                                                            NCBI_TaxID-11111;
```

1 Similarity 100. 28; Conservative

Best Local Matches

ö

Gaps

ö

Indels

0

0; Mismatches

```
NCBI_TaxID=31644;
                                                                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                 POLG_HCVHK
Q01403;
 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                       POLG_HCVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                        .
G
                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPTIDASE.
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTEWIAL).
MATRIX PROTEIN (POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIAL).
                                                                                                                                                                                                                                                                                                                       MEDLINE—93019030; PubMed-1383400;
Abe K., Inchauspe G., Fujisawa K.;
"Genomic characterization and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN A AND GIXCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                        01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (Isolate HCV-476) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                               ;
                        Length 513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Coat protein; Envelope protein;
                                               Indels
55704 MW; 943F31E3514CDEF3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GICNAC. . .)
                      Watch 94.2%; Score 145; DB 1; LA Local Similarity 96.4%; Pred. No. 4.1e-14; es 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BY SIMILARITY
N-LINKED (GLC
                                                                     1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transmembrane; Nonstructural protein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D10688; BAA01530.1; -.
InterPro; IPR002522; HCV_caps1d.
InterPro; IPR002513; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
InterPro; IPR002531; HCV_env.
IPfam; PP01543; HCV_caps2d; I.
Pfam; PF01543; HCV_caps2d; I.
Pfam; PF01539; HCV_core; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Co
                                                                                                                                                     STANDARD;
513 AA;
                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                      NCBI_TaxID=31643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01560;
                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                     POLG_HCVH4
 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                        Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INIT_MET
                                                                                                                                                                 001404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                              RESULT 3
POLG_HCVH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                              Matches
 So
                                                                                          g
                                                                                                                                                      DA PP PP PP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð,
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifiated and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE-9310930; PubMed=1383400;
Abe K., Inchauspe G., Fujisawa K.;
Abe K., Inchauspe G., Fujisawa K.;
Genomic characterization and mutation rate of hepatitis C virus
isolated from a patient who contracted hepatitis during an epidemic
of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2739(1992).
I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).

Hepatitis C virus (Isolate HCV-KF) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ransmembrane; Nonstructural protein.
INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                           ;
                                                                                         Score 145; DB 1; Length 520; Pred. No. 4.1e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CAPSID PROTEIN C (POTENTIAL). MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                           1; Indels
520 520
520 AA; 56499 MW; AA135246CF20D525 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY SIMILARITY.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                       520 AA
                                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                        InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_core.
InterPro; IPR00253; HCV_core.
InterPro; IPR002531; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01549; HCV_core; 1.
Pfam; PF01560; HCV_env; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D10687; BAA01529.1; -. PIR; JQ1925; JQ1925.
                                                                                      Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
```

```
ď
                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the SAMES Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEDLINE—92230232; PubMed=1314459;

MEDLINE—92230232; PubMed=1314459;

MEDLINE—92230232; PubMed=1314459;

MEDLINE—92230232; PubMed=1314459;

A Tanaka T., Fukuda S., Tsuda F., Mishiro S.;

Tanaka T., Fukuda S., Tsuda F., Mishiro S.;

Tenaka T., Fukuda S., Tsuda F., Mishiro S.;

Tranaka T., Tsukuda S., Tsuda F., Mishiro S., Tsukuda T., Tsukuda T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ERMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (FOTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix protein (Envelope protein M); Major envelope protein E, Nonstructural proteins NS1 and NS2] (Fragment).
Hepatitis C virus (isolate HC-J5) (HCV).
Hepatitus:
                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   K
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                           ö
                                                                                                                                                 Score 145; DB 1; Length 520;
Pred. No. 4.1e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coat protein; Envelope protein;
                                               449 N-LINKED (GLCNAC. . .) (PC 520 5476 MW; 1D2BD0A6FF27349B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                     737 AA
                                                                                                                                                                                                                                                   Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR00252; HCV_capsid.
Interpro; IPR00252; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002531; HCV_env.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01549; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Probom; PD186062; HCV_NS1; 1.
POlyprotein; Glycoprotein; Co: Transmembrane; Nonstructing Co: INTP_MET
                                                                                                                                                 Ouery Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
424
431
449
520
520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116
192
384
734
                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVJ5
ID POLG_HCVJ5
AC P27960;
PT 01-AUG-1992
                                                  CARBOHYD
NON_TER
SEQUENCE
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                  RESULT
     SFFF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                   g
```

POTENTIAL.  N-LINKED (GLCNAC) (POTENTIAL).	94.2%; Score 145; DB 1; Length 737; ; Conservative 0; Mismatches 1; Indels 0; Gaps 0; PRPORKTRRNTYRRODVKFPGGGIVG 28 	STANDARD: PRT; 737 AA.  1. 23, Last sequence update)  1. 40, Last annotation update)  1. 40, Last annotation update)  2. Ann [Contains: Cappsid protein C (Core protein); Matrix pe protein M); Major envelope protein E; Nonstructural NS2] (Fragment).  2. (Isolate HC-J7) (HCV).  2. (Solate HC-J7) (HCV).	SEQUENCE FROM N.A.  MEDLINE-9223023; PubMed-1314459;  MEDLINE-92230232; PubMed-1314459;  A Tanaka T., Kural K., Okada S.I., Yamamoto K., Lizuka H.,  Tanaka T., Fukuda S., Tsuda F., Mishiro S.;  Tanaka T., Fukuda S., Tsuda F., Mishiro S.;  "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";  Virology 188:331-341(1992).  L Virology 188:331-341(1992).  L FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC. SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  SJUBUNT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:  PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).  EMBL; D10077; BAA00971.1;  InterPro; IPR002522; HCV_core.  InterPro; IPR002521; HCV_core.  InterPro; IPR002519; HCV_eore.  InterPro; IPR002531; HCV_eore.  InterPro; IPR002531; HCV_eore.  InterPro; IPR002531; HCV_eore.
369 1196 2209 2209 305 4417 448 448 448 554 558 627 81207	94.2% 96.4% vative NTYRRPQD	STANDARD: 1. 23, Crea 1. 23, Last 1. 40, Last 1. 40, Last 2. 40, Last 2. 40, Last 3. 40, Last 3. 40, Last 4.	PubMed-1 7., Okad 7., Okad 7., Tsud 10.e of a 11.(1992) 11.(1992) 11.(1992) 12.ELOP 12.ELOPE 12.ELOPE 13.ELYCOPRO	Institutions in the control of the c
347 196 2209 2209 2304 417 417 423 430 448 448 477 534 558 578 627 627 737 AA;	ilarity Conserv PORKTKR	7 STAN 92 (Rel. 2 92 (Rel. 4 01 (Rel. 4 11yprotein Envelope E NSI and NS C virus ( SSRNA posius)	30232; Frustal Russell	PROT ent e Swiss n Bioinf d this s quires a quires a PR002522 PR002521 PR002531
TRANSMEM CARBOHYD CAR	Query Match Best Local Simil. Matches 27; C 1 PKPQ   1	SULT 6 POLG HVV7 POLG HVV7 P27961, 01-A0G-1992 (Rel. 01-A0G-1992 (Rel. 01-A0G-1992 (Rel. 01-A0G-1992 (Rel. DOCT-2001 (Rel. Genome polyprotein Genome polyprotein Protein (Envelope Protein SI and N Protein SI and	SEQUENCE FROM MEDLINE-923302 Okamoto H., Ku Tanaka T., Fuku Full-length S. homology 188:3 virology 188:3	This SWISS- between the Europea use by no modified an entities re or send an entities bloor interPro; I Pfam, PF015
	A Ma Ma O Y	RESULT POLCE IN THE POLCE IN TH	# # # # # # # # # # # # # # # # # # #	30000000000000000000000000000000000000

S

Fri Aug

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 D1-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein [P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3-4-22.-); Proteases/helicase NS3 (P70) (Heparivith)
(EC 3-4-21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate E1 avivinae no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                               CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.
REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-96235224; PubMed*8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 737;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 145; DB 1; Length 73
Pred. No. 6.1e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N.LINKED
N.L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 Similarity 96.4 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      299
305
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      737
737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
POLG_HCVBK
                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                 DR NO DOR NO DOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FT FT FT SO SO SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 5 E
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                         MEDLINE-9827386; PubMed-9568891;

X MEDLINE-9827386; PubMed-9568891;

A StellKuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;

T "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C. Tomelex of NS3 protease and NS4A peptide of BK strain hepatitis C. T. virus: a Z.2-A resolution structure in a hexagonal crystal form.";

T virus: a Z.2-A resolution structure in a hexagonal crystal form.";

T VIRONON: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAX PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral postition, Commonly with Asp or Glu in the P6 position. Cys or Thr in Pl and Ser or Ala in Pl'.

C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
SUBUNTT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                    ø
                                                                             LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
The crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
phosphorylation mediated by CAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; PubMed-8861916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
543; HCV_capsid; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1A10; 25-MAR-98.
1JXP; 14-JAN-98.
1NS3; 08-APR-98.
1C2P; 15-NOV-00.
1CSJ; 08-NOV-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1GX5; 09-APR-02.
1GX6; 10-APR-02.
1QUV; 26-JUN-00.
8OHM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEROPS; S29.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01539;
PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                       -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PDB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB
```

```
SEEEES
                                                                   à
                                                          CELLULAR AMINOPEPTIDASE.
CAPESID PROTEIN C (POTEWTIAL).
MAJOR ENVELOPE PROTEIN (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS1/EZ (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTEWTIAL).
PROTEASE/HELICASE NS3 (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTEWTIAL).
RNA-DIRECTED RNA POLYMERASE (POTEWTIAL).
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF001506; HCV_NS5a; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Cat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                  (POTENTIAL).
                                                                                                                                                                                                              (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                     POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
                                                                                                                                                        N-LINKED (GLCNAC.
                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                  DECH BOX.
                                                                                                                                                                                                                                            N-LINKED
                                                                                                                     369
111083
11107
111083
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
11109
                                                                                                                                                                                                                                                                                                                      085
087
092
094
097
103
                                                                                                                                                                                                                                                                                                                      10082
10086
10090
10090
111009
111120
11120
11130
11130
11168
11168
11168
                                                                                                                                                  1316
196
209
234
250
305
417
                                                                                                                                                                                                                                                            25240
2529
2788
2788
1031
1039
1050
1059
1059
                                                                                                                                                                                                   430
                                                3D-structure, INIT_MET
                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                      ACT_SITE
NP_BIND
                                                                                                                          ACT_SITE
ACT_SITE
                                                                                                                                                                                            CARBOHYD
                                                                                                                                                         CARBOHYD
                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                               CARBOHYD
                                                                                                                      TRANSMEM
                                                                                                                                                               CARBOHYD
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                 STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                               STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                            STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                   CHAIN
CHAIN
CHAIN
                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                       HELIX
                                                                  CHAIN
                                                                        CHAIN
                                                                                                     CHAIN
                                                                                                           CHAIN
                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                    HELIX
                                                                                                                                                                                                                                                                                                           TURN
                                                                                                                                                                                                                                                                                                                            TURN
                                                                                                                                                                                                                                                                                                                                                         TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (N81); Protein P2 (Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS48 (P56); Nonstructural protein NS58 (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
MS5B (P66) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate Japanese) (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            [RNA][N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSIGNS OF TWO PROTEIN S: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDILINE-91088550; PubMed-2175903;
Kato N., Hilikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molacular cloning of the human hepatitis C virus genome from
"Japanese patients with non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990);
                                                                                                                                                             Length 3010;
                                                                                                                                                                                                            Indels
                                                                                                               327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                             Score 145; DB 1;
Pred. No. 2.9e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3010 AA
                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                             94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D90208; BAA14233.1; -.
                                                                                                                                                         Query Match 94.2
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISCUSSION OF SEQUENCE.
1188
1197
1202
1204
1688
                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A39253; GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCB1_TaxID=11116;
11187
11189
11203
1680
                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVJA
P26662;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                           HELIX
TURN
STRAND
SEQUENCE
TURN
                                                                                                                                                                                                                                                                                                                                                                         RESULT 8
POLG_HCVJA
```

```
Transmembrane; Nostructural Fransferase; RNA-directed RNA polymerase;
W Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Helicase; ATP-binding;
TRMAVED FROM CAPSID PROTEIN C BY THE
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
TO CHAIN 1007 1615 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
TO CHAIN 1007 1615 PROTEINGARE, NSI (POTENTIAL).
TO CHAIN 1616 1863 2013 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
TO CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
TO CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
TO CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
TO CHAIN 1007 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -LINKED
                                                                                                                                                                                                                                                                                 InterPro; IPR001650; Helicase.
InterPro; IPR001650; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR001095; RNA_pol_DS_PI.
Interpro; IPR001505; RNA_pol_DS_PI.
Interpro; IPR01505; RNA_PS_PI.
INTERPROPERTY.
INTER
                                                                                                                                                                                                                                                 HCV_NS5a.
HCV_RdRP.
Helicase_C
                                                                                                      HCV_core.
HCV_env.
HCV_NS1.
                                                                                                                                                                HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                       InterPro; IPR001410; D
InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002519; H
InterPro; IPR002519; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1165
                                                                                                                                                                                                                                             InterPro; IPR002868;
InterPro; IPR002166;
InterPro; IPR001650;
                                                                                                                                                                                   InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                              InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥.
                 MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SITE
```

Gaps

ö

Length 3010;

Score 145; DB 1; Length 30 Pred. No. 2.9e-13; Mismatches 1; Indels

ö

Conservative

27;

Similarity

Query Match Best Local S Matches 27

94.2%; 96.4%;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Usage by and for commercial
                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
66-00me polyprotein [Contains: Capsid protein C (Core protein) (P22);
670-69 (GPV6) (MSI); Protein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (MSI); Protein E7; Nonstructural protein NS2 (P21)
6EC 3.4.29-7); Protease/helicase NS3 (P70) (Hepacivirin)
6EC 3.4.21-98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); Nonstructural 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kim J.L., Morganstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko M.A., Lin C., Caron P.R.; "Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           H.V., Weber P.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ε
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ВУ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SUBJUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED B LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-!- PTM: THE STRUCTURAL PROTEINS C, EI AND E2 ARE PRODUCED BY PROTECLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-!- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
-!- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              use by non-profit institutions as long as its content modified and this statement is not removed. Usage by an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V.
"Structure of the hepatitis C virus RNA helicase domain."
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
                                                                                                                                                                                            PRT; 3011 AA
                               PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-92052256; Pubmed-1658800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-98154321; PubMed-9493270;
                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  of unwinding.";
Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA)(N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Prince A.M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                            POLG_HCVH
Н
                                                                                                                                                                                                                        g
```

```
Gaps
   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1627
1636
1652
A: 327142 MW; 772CBB29CCD94753 CRC64;
     ??
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 145; DB 1; 1
Pred. No. 2.9e-13;
0; Mismatches 1;
   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
 1584 159
1598 159
1606 161
1614 1612
1622 1622
1637 163
1640 165
3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                        11376
11378
11382
11389
11389
11410
11419
11432
11438
11450
11480
11480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVJ6
ID POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HELIX
TURND
STRAND
STRAND
STRAND
HELIX
HELIX
HELIX
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
HELIX
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
 CARBOHYD
CARBOHYD
                                 RESULT 10
   ò
 (See http://www.isb-sib.ch/announce/
              EMBL. M67463; AAA4534.1; -..

BRIE. M67463; AAA4535.1; -..

BRIE. M67463; AAA46351; HCV. Core.

BRIE. M67463; HCV. ASA.

BRIE. M67463; HCV. NSA.

BRIE. M67463; HCV. NSA.

BRIE. M67463; HCV. ASA.

BRIE. M674641; HCV. ASA.

BRIE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINDEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          titles requires a license agreement (
  send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
CCHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
CARBOHYD
C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
THE FEET TENT TO THE FEET THE
```

```
Pfam; PF01506; HCV_MS5a; 1.

Pfam; PF00391; Viral_RARP; 1.

Promom; PD186062; HCV_MS1; 1.

R Probom; PD186062; HCV_MS1; 1.

R SMART; SM00487; DEXDC; 1.

R Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

T CHAIN 115 CAPSID PROTEIN C BY THE CELLULAR ANIOPEPTIDASE.

T CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).

T CHAIN 192 383 MADOR ENVELOPE PROTEIN E.

T CHAIN 384 733 MADOR ENVELOPE PROTEIN NSI (POTENTIAL).

T CHAIN 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1619 PROTEIN NSI (POTENTIAL).

T CHAIN 1619 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

T CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).

T CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).

T CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).

THAN 2018 3033 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-406-1992 (Rel. 23, Created)
01-406-1992 (Rel. 23, Last sequence update)
01-406-1992 (Rel. 21, Last sequence update)
01-606-1992 (Rel. 41, Last annotation update)
03-8-FBB-2003 (Rel. 41, Last annotation update)
04-606 (Rel. 41, Last annotation update)
05-606 (GP60) (Rel. 41, Protein El (GP32) (GP35); Envelope glycoprotein E2
05-606 (GP06) (NS1); Protein El (GP32) (GP35); Envelope glycoprotein E2
05-606 (GP10) (NS1); Protein Protein NS4 (P30) (Rel. 22.-); Protein Reservation (Rel. 41, 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
CHARGE RELAY SYSTEM ()
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F957F5C1A273BE9E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 145; DB 1;
Pred. No. 2.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Sim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJ8

1D POLG_H
AC POLG_H
DT 01-AUG
DT 01-AUG
DT 28-FEB
DE GENORIO
DE (GF8)
DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=92044440; PubMed=1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Adhida A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
and divergent regions. "In the reported isolates for conserved
and divergent regions." The SAGIL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presentsor polyprotein, commonly with Asp or Glu in the P6
precursor polyprotein, commonly with an pil.
C -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                           01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N81); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NSA (P49; Nonstructural protein NSA (P47); Nonstructural protein NSA (P46); Nonstructural protein NSA (P46); Nonstructural protein NSA (P46); Nonstructural protein NSA (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)}.
Hepatitis C virus (isolate HC-J6) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBUNTT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR00166; HVV_Rdae.C. InterPro; IPR001650; Helicase_C. InterPro; IPR007095; RNA_pol_DS_PS. InterPro; IPR007094; RNA_pol_DS_PS. Pfam; PF01543; HCV_capsid; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01542; HVC_COFE; I
PF01539; HCV_EN; 1.
PF01560; HCV_NS1; 1.
PF01538; HCV_NS2; 1.
PF02507; HCV_NS3; 1.
PF01006; HCV_NS4s; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D00944; BAA00792.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR, JQ1303, JQ1303.
HSSP, P27958, 1HEI.
MEROPS, S29.001; -.
MEROPS, U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
```

ö

Gaps

ö

Pfam; Pfam;

Pfam; Pfam; Pfam; Pfam;

```
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                   N-LINKED
N-L
                                                                                                                              N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                              DECH BOX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best_Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVTW
P29846;
                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
     CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses;
                                                                                                                                CARBOHYD
                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                       between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAPSID PROTEIN C (POTENTIAL).
MACHRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS46 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
                                                                                                     SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
Full-length sequence of a hepatitis C virus genome having poor
homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core brotein; Coat protein; Envelope protein; Halicase; APT-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease: INIT_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                   SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR002521; HCV_core.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_RNS.
InterPro; IPR00219; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001496; HCV_NS5a.
InterPro; IPR002166; HCV_RSAP.
InterPro; IPR002166; HCV_RNSP.
InterPro; IPR007095; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR007094; RNA_POI_DS_PS.
InterPro; IPR01643; HCV_CORPII,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8. HSSP; P27981, 1HEI. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF01506; HCV_NSSa; 1
Pfam; PF00998; Viral_RdRP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01539;
PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       116
192
384
734
1011
1620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last amoutation update)
15-SEP-2003 (Rel. 42, Last amoutation update)
16-GPG93 (Rel. 42, Last amoutation update)
Envelope polyprotein [Contains: Capsid protein C (Core protein E2 (GPG9) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivitin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P50); Nonstructural protein NS5A (P60) (P70); Nonstructural protein NS5A (P60) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                    SIMILARITY).
SIMILARITY).
                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                        (POTENTIAL)
                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                          MW; 1A173E7E3381FD1A CRC64;
 (BY
(BY
(BY
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 145; DB 1;
Pred. No. 2.9e-13;
                                                                                     (GLCNAC.
                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                          (GLCNAC
                                                                                                                                            (GLCNAC
                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                      GLCNAC
                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                      ATP (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-SEP-2003 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last sequence update)
16-SEP-2003 (Rel. 42, Last sequence update)
16-SEP-2003 (Rel. 42, Last sequence)
16-SEP-2003 (Rel. 42, Last s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-91172826; PubMed-1848704;
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
Bradley D.W., Kuo G., Houghton M.;
"Genetic organization and diversity of the hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 88:2461-2455(1991).
-!-FUNCTION: THE SMALL PROTEINS NSZA, 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC. . .) (POTENTIAL)
WW; AAD267D55CDFE215 CRC64;
                                                POTENTIAL
                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 138; DB 1;
Pred. No. 3.3e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d. No. 3.36
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6 KPORKTKRNTNRRPODVKFPGGGOIVG 32
        N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 KPORKTKRNTYRRPODVKFPGGGGIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89.68;
96.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M62321; AAA45676.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                         2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11104;
                                                                                                                                                                                                                                            623
645
2041
2077
2240
25529
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ma.
Local Sin.
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCV1
P26664;
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                         CARBOHYD
                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCV1
            25.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DATE THE RESERVENCE OF THE PARKET AND A COUNTY O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPTIDASE.

CORE PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN NSJ (EO POTENTIAL).

PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ресн вох.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
R InterPro; IPR004109; HCV_NS3.
R InterPro; IPR004109; HCV_NS4.
InterPro; IPR001499; HCV_NS5a.
InterPro; IPR001669; HCV_NS5a.
InterPro; IPR002669; HCV_NS5a.
InterPro; IPR007095; RNA_POl_DS_PS.
InterPro; IPR007095; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01538; HCV_NS2; 1.
PF02907; HCV_NS3; 1.
PF01006; HCV_NS4s; 1.
PF01001; HCV_NS4b; 1.
PF01506; HCV_NS5s; 1.
PF00271; helicase_C; 1.
PF00998; Viral_RdRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01539; HCV_env;
PF01560; HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1006
11615
22013
30103
30103
10083
11165
11165
1213
2233
4117
4117
4117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProDom;
SMART; SI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
```

```
MEDLINE-9232543; PubMed-1629222;
Leduc R., Molloy S.S., Thorne B.A., Thomas G.;
Leduc R., Molloy S.S., Thorne B.A., Thomas G.;
Leduc R., Molloy S.S., Thorne B.A., Thomas G.;
Activation of human furin precursor processing endoprotease occurs by an intramolecular autoprotechytic cleavage.";
J. Biol. Chem. 267:14304-14308(1992).
-!-FUNCTION: FUBLIN IS LIKELY TO REPRESENT THE UBIQUITOUS ENDOPROTEASE ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF CLEAVAGE AT THE RX(K/K)R CONSENSUS MOTIF.
-!-CATALYTIC ACTIVITY: Release of mature proteins from their proproteins by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa bonds any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and von Willebrand factor from their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 296-794 FROM N.A. MEDIAGE SEQUENCE OF 296-794 FROM N.A. MEDIATE-8705388; PubMed-3023061; Reobrock A.J.M., Schalken J.A., Leunissen J.A.M., Onnekink C., Bloemers H.P.J., van de Ven W.J.M.; "Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein."; EMBO J. 5:2197-2202(1986).
                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
MEDLINE-94291619; PubMed-8020465;
Siezen R.J., Creemers J.W.M., van de Ven W.J.M.;
"Homology modelling of the catalytic domain of human furin. A model
for the eukaryotic subtilisin-like proprotein convertases.";
Eur. J. Blochem. 222:255-266(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "CDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues."; DNA Cell Biol. 10:319-328(1991).
                                                                                                                                                                                                                                                                                                                              Homo saplens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-90175002; PubMed-2408021; van den Ouweland A.M.W., van Duijnhoven H.L.P., Keizer G.D., van den Ouweland A.M.W., van de Ven W.J.M.; "Structural homology between the human fur gene product and the subtilisin-like protease encoded by yeast KEX2."; Nucleic Acids Res. 18:664-664(1990).
                                                                                                                                                                                                                                                                                       acid residue
                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-91321735, Pubmed-1713771;
Barr P.J., Mason O.B., Landsberg K.E., Wong P.A., Kiefer M.C.,
                                   Length 3011;
                      ; DB 1; Lens. 4 6e-12; Indels
   327197 MW; 65F8C9447FCESAF9 CRC64;
                                                                                                                                                                                                   PURIL HUMAN STANDARD; PRT; 794 AA. P09958; 014336; 01-MAR-1989 (Rel. 10, Created) 01-MAR-1990 (Rel. 14, Last sequence update) 15-SEP-2003 (Rel. 42, Last annotation update) Furin precursor (EC 3.4.21.75) (Paired basic amino acclearing enzyme) (PACE) (Dibasic processing enzyme).
                                                  Pred. No. 4.66
1; Mismatches
                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                             Score 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COFACTOR: CALCIUM-DEPENDENT
                                   89.08;
                                                  llarity 89.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          respective precursors.
    3011 AA;
                                               Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE-Blood;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Brake A.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROCESSING
    SEQUENCE
                                   Query Match
                                                                                                                                                                                          FURI_HUMAN
                                                                                                                                                                         RESULT 14
                                                                                                                       g
                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN NSI/EZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/EZ (POTENTIAL).
PROYEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                              InterPro; IPR002166; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007094; RNA_pol_DS_ES.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                      V_capsid.
                                                                                                                                     HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4a.
                                                                                                          HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01006; HCV_NS4a; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01050; HCV_NS5a; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00998; Viral_RdFR; 1. ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          729
1006
11615
1862
2013
3011
369
1107
11165
11165
1237
1319
1319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234
305
417
423
430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART; SM00487; DEXDC;
                                                                                                                                                                                                                  IPR002868; H
IPR002166; H
              PDB; 1A1V; 16-FEB-99,
PDB; 1HEI; 25-NOV-98,
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                           InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                      InterPro; IPR002521;
                                                                                                                        InterPro; IPR002519;
A39166; GNWVC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      347
11083
11107
11165
11230
1316
1316
209
234
417
417
4233
                                                                                                                                                                                                                                                                                                                                                                          PF02907; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                    InterPro;
                                                                                                                                                                                     InterPro;
                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                              Pfam;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation - the European Bioinformatics. The Three are no restrictions on its use by non-profit institutions as long as its content is in no way modifiated and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                  SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE TRANS GOLGI NETWORK. PROPEPTIDE CLEAVAGE IS A PREREQUISITE FOR EXIT OF FURIN MOLECULES OUT OF THE ENDOPLASMIC RETICULUM (ER). SECOND CLEAVAGE IN THE PROPEPTIDE OCCUR IN THE TRANS GOLGI NETWORK FURN), WHICH IS FOLLOWED BY THE RELEASE OF THE PROPEPTIDE BOUND TO FURIN AND THE ACTIVATION OF FURIN.
                                                                                                                                                                                                                            TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED UBIQUITOUSLY.
DOMAIN: CONTAINS A CYTOPLASMIC DOMAIN RESPONSIBLE FOR ITS TGN
LOCALIZATION AND RECYCLING FROM THE CELL SURRACE.
PTW: THE PROPERTIDE IS AUTOCATALYTICALLY REMOVED THROUGH AN
INTRAMOLECULAR CLEAVAGE PROBABLY IN THE ENDOPLASMIC RETICULUM
(ER). IN THE TGN THE SECOND CLEAVAGE IN THE PROPEPTIDE COULD LEAD
TO THE ACTIVATION OF FURIN.
SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
ENZYME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS00136; SUBTILASE_ASP; 1.
PROSITE; PS00137; SUBTILASE_HIS; 1.
PROSITE; PS00138; SUBTILASE_SER; 1.
Hydrolase; Serine protease; Transmembrane; Glycoprotein; Signal; Zymogen; Calcium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLEAVAGE (SECOND AUTO').
CLEAVAGE (FIRST AUTO').
CLELL ATTACHMENT SITE (POTENTIAL).
CELL SURFACE SIGNAL.
TRANS GOLGI NETWORK SIGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10C44DD5892EF85D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      136950, -. Golgi apparatus, TAS. GO:00047794; C:Golgi apparatus, TAS. GO:0004276; F:furin activity; TAS. GO:0007267; P:cell-cell signaling; TAS. GO:0005508; P:proteolysis and peptidolysis; TAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPPR006213; Furin_repeat.
InterPro; IPR002844; P_domain.
InterPro: IPR002039; Peptidase_S8.
Pfam; PF01483; P_proprotein; PARTIAL.
Pfam; PF00082; Peptidase_S8; 1.
PRINTS; PR00723; SUBTILISIN.
ProDom; PD000717; P_domain; 1.
SNART; SM00201; F0; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FURIN.
CYS-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  86678 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X17094; CAA34948.1; -. EMBL; X04329; CAA27860.1; -. EMBL; A06939; CAA00605.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S08.071; -. Genew; HGNC:8568; FURIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    107
500
762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  794 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A39552; KXHUF.
HSSP; Q99405; 1MPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
   <u>.</u>
                                                           -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  888
88888888888888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE STATES STA
```

DB 1; Length 794;

Score 60; DB 1, Pred. No. 0.41;

39.0%;

Best Local Similarity

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SUBCELLULAR LOCATION: SEEMS TO BE LOCALIZED INTRACELLULARLY TO THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              complement component C3 and von Willebrand factor from their respective precursors.

COFACTOR: CALCIUM-DEPENDENT (BY SIMILARITY).

BN2YME REGULATION: COULD BE INHIBITED BY THE NOT SECONDLY CLEAVED
  Gaps
                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY. SIMILARITY: Contains 1 homo B/P domain.
                                                                                                                                              01-NOV-1991 (Rel. 20, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Furin precursor (EC 34.21.75) (Paired basic amino acid residue cleaving enzyme) (PACE) (Dibasic processing enzyme).
 ö
                                                                                                                                                                                                                                                                                                                       "Structure and expression of mouse furin, a yeast Kex2-related protease. Lack of processing of coexpressed prorenin in GH4Cl
                                                                                                                                                                                                                                                                                               Hatsuzawa K., Hosaka M., Nakagawa T., Nagase M., Shoda A.,
 Indels
  4
                                                                                                              793 AA
  Mismatches
                                                                                                                                                                                                                                                                                                                                                            J. Biol. Chem. 265:22075-22078(1990).
                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
MEDLINE-91093035; Pubmed-2266110;
                                     :|:|||: |: | | |||
102 KRRTKRDVYQEPTDPKFP 119
  .,
4
                        21
                                                                                                                                    01-NOV-1991 (Rel. 20, Created)
                       4 ORKTKRNTYRRPODVKFP
                                                                                                                                                                                                                                                                                                           Murakami K., Nakayama K.;
 10; Conservative
                                                                                                              STANDARD;
                                                                                                                                                                                    cleaving enzyme) (PAC FURIN OR FUR OR PCSK3
                                                                                                                                                                                                             musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Liver
                                                                                                              MOUSE
                                                                                                                                                                                                                                                                                                                                                  cells.
Matches
                                                                                                                                                                                                             Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  +
                                                                                     RESULT
                                               g
                                                                                                              à
```

```
PIR; A23679; KANST.

R HSSP; Q9405; IMPT.

R HSSP; Q9405; IMPT.

R HSRDPS; SO80.071; Furin_repeat.

R InterPro; IPR000284; P_domain.

InterPro; IPR000284; P_domain.

R InterPro; IPR00029; Peptidase_S8.

R Pfam; PR001483; P_DP0TOTCtein; PARTIAL.

R PRNOS; PR00029; POSTIGASE_S8; I.

R PRNTS; PR00723; SUBTILISSIN.

R PROSTIE; PS00136; SUBTILIASE_ASP; I.

R PROSTIE; PS00136; SUBTILIASE_HIS; I.

R PROSTIE; PS00138; SUBTILIASE_HIS; I.

R PROSTIE; PS00138; SUBTILIASE_ESR; I.

R HYdrolase; Serine protease; Transmembrane; Glycoprotein; Signal; M Zymogen; Calcium.
                                                                                                                                                                                                                                                                                                                                       BY SIMILARITY.
FURIN.
CYS-RCH.
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
POTENTIAL.
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
CLEAVAGE (SECOND AUTO-).
CLEAVAGE (SECOND AUTO-).
CLEAVAGE (SIRST AUTO-).
CLEAVAGE (SIRST AUTO-).
CLEAVAGE (SIRST AUTO-).
CELL SURFACE SIGNAL.
TRANS GOLGI NETWORK SYGNAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 35.7%; Score 55; DB 1; Length 793; Best Local Similarity 50.0%; Pred. No. 2.3; Matches 9; Conservative 4; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86804 MW;
EMBL; X54056; CAA37988.1; -. EMBL; L26489; AAA37643.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          793 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
DISULFID
DISULFID
CARBOHYD
CARBOHYD
CARBOHYD
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
ACT_SITE
ACT_SITE
```

.rch completed: August 7, 2003, 11:19:58 .o time : 5.90909 secs

4 ORKTKRNTYRRPODVKFP 21

qq

ò

```
August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                         830525
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                   tal number of hits satisfying chosen parameters:
                                                                                                                                                                US-09-491-146A-25
154
1 PKPQRKTKRNTYRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                       830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_fung1:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0 Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sb_mhc:*
                                                                                                                                                                   Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database :
                                                                                                                                                                                                                                                                                       Searched:
                                                                                                  Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	P90364 hepatitis c	Q81831 hepatitis c	_	Q68575 hepatitis c	_			Q81807 hepatitis c					_	_	Q8qp85 hepatitis c	Q8qp87 hepatitis c
SUMMARIES	Ω	P90364	2 Q81831	PENXAGO :	2 Q68575	2 Q91AU0			9 081807							•	2 Q8QP87
	% Query Match Length DB	105 12	106 12	125 12	191 12	3010 12	•	. ,	109 12	•	45 12	45 12	60 12	61 12	73 12	100 12	100 12
	% Ouery Match	100.0	100.0	100.0	100.0	100.0	100.0	100.0	97.4	94.8	94.8	94.8	94.8	94.8	94.8	94.8	94.8
	Score	154	154	154	154	154	154	154	150	146	146	146	146	146	146	146	146
	Result No.	-	7	e	4	2	9	7	œ	σ	10	11	12	13	14	15	16

7 146 94.8 100 12 080P72 08qP72 hepatitis 8 146 94.8 100 12 080P71 08QP74 08qP71 hepatitis 9 146 94.8 100 12 080P74 08qP74 08qP77 hepatitis 0 146 94.8 100 12 080P80 08qP80 hepatitis 1 146 94.8 100 12 080P80 08qP80 hepatitis	7. 22 146 94.8 100 12 080P84 089P84 080P84 hepatitis c 23 146 94.8 100 12 080P78 080P78 080P78 hepatitis c 24 146 94.8 100 12 080P79 080P75 hepatitis c 25 146 94.8 100 12 080P81 080P91 080P91 hepatitis c 27 146 94.8 100 12 080P81 080P81 080P81 115 c	8 146 94.8 100 12 080P76 089P76 hepatitis 9 146 94.8 100 12 080P86 089P86 089P86 hepatitis 0 146 94.8 109 12 081340 081340 hepatitis 1 146 94.8 114 12 080M14 ORGANIA DESCRIPTION	2 146 94.8 114 12 068892 Q68892 hepatitis 3 146 94.8 114 12 068893 Q68893 hepatitis 4 146 94.8 119 12 Q8BCX2 Q8bcx2 hepatitis 5 146 94.8 119 12 Q8BCX2 Q8bcx2 hepatitis 5 146 94.8 119 12 Q8BCX2	6 146 94.8 119 12 QBBCXO QBbcxO hepatitis 7 146 94.8 119 12 QBBCW9 QBbcw9 hepatitis 8 146 94.8 119 12 QBBCW7 QBbcw9 hepatitis 9 146 94.8 119 12 QBBCW7 QBbcw7 hepatitis	0 146 94.8 119 12 QBBCW6 QBbcw6 hepatitis 1 146 94.8 119 12 QBBCW4 QBbcwW hepatitis 2 146 94.8 112 12 QBQRJ1 QBqTJ1 hepatitis 3 146 94.8 122 12 QBQRJ6 QBqml6 hepatitis	5 146 94.8 122 12 080R34 080R74 hepatitis ALICOMENTS	RESULT 1 P90364 ID P90364 PRELIMINARY; PRT; 105 AA.	0364; -MAY-1997 (TrEMBLrel. 03, Created) -MAY-1997 (TrEMBLrel. 03, Last segment)	-OCT-2002 (TEMBLE) 22, Last annotation update) re protein (Genome polyprotein) (Fragment).	<pre>C Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; C Hepacivirus. X NCBI_TaxID=11103; N fil</pre>	SEQUENCE FROM N.A. A Ohno T., Mizokami M.; T "Determination of nine genotypes of hepatitis C virus using PCR	1 mechod. L Thesis (1995), Nagoya City University Medical School. N [2]	QUENCE FROM N.A. :DLINE-96305314; PubMed-8712927; no T., Mizokami M., Saleh M.G., Orito E., Obba KI., Wu R.	oide T., Tibbs C.J., Nouri-Aria K.T., Tokudome S., Williams R., sefulness and limitation of phylogenetic analysis for hepatit: rus core region: application to isolates from Egiptian and Yer	RY patients."; RL Arch. virol. 1113(1996). CC -: SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A CC -: SUBUNIT: THE VIRION OF THIS VIRUS OF TWO PROTEINS. CC PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF CC PROTEIN G. AND MRNA (BY SIMILARITY).	R EMBL; D82034; BAA11519.1; R InterPro; JPR002522; HCV_capsid. R Pfam; PP01543; HCV_capsid; 1. W Polyprotein.	T NON_TER 1 1 1 . T NON_TER 105 105 Q SEQUENCE 105 AA; 11870 MW; C7BA40B284025A49 CRC64;
								~ 11 1				- 144	_ = = =				_ m 0,

Gaps

ö

Indels

Length 125;

```
"Molecular cloning and expression of hepatitis C virus core protein and production of monoclonal antibodies to the recombinant protein.", Asian Pac. J. Allergy immunol. 14:0-0(1996).

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; U55284; AAB00215.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HCV-S1; MEDLINB-21440119; PubMed-11556407; Lim S.G., Tan Y.H.; Lim S.P., Khu Y.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.; Lim S.P., Khu W.L., Hong W.J., Tay A., Ting A.E., Lim S.G., Tan Y.H.; Indentification and molecular characterization of the complete genome of a Singapore isolate of hepatitis C virus: sequence comparison with other strains and phylogenetic analysis.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Songsivilai S., Dharakul T., Kunkitti R., Thepthai C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191
20840 MW; 4AAE63444D8329E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                       Score 154; DB 12;
Pred. No. 1.2e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 154; DB 12;
Pred. No. 1.9e-15;
; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                              01-07-1996 (TrEMBLrel. 01, Created)
01-NOY-1996 (TrEMBLrel. 01, Last sequence updi
01-07-2002 (TrEMBLrel. 22, Last annotation updi
02-07-protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3010 AA
                                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0%; Prea. ...
                                                                                                                                                                                       1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPORKTKRNTYRRPODVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam: PF01543; HCV_capsid; 1.
Pfam: PF01542; HCV_core; 1.
                                    100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                U1-DEC-2001 (TEMBLE). 19, 01-DEC-2001 (TEMBLE). 19, 01-MAR-2003 (TEMBLE). 23, Genome polyprotein.
                                                                Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best_Local Similarity 100.0
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. STRAIN-BB51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
NON_TER 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                             Q68575
Q68575;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           091AU0
                                                                                                                                                                                                                                                                                                     RESULT 4
Q68575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
Q91AU0
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                DDT BERNAR BERNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus,
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
E2/NS1 protein (Genome polyprotein) (Fragment).
Hepatitis C virus type 2.
Hepatitis C virus type 2.
Hepatitis C virus type 2.
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEGCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GINCORROTEIN E. THE NUCLEGCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-94351179; Pubmed-7520922;
Nakazawa T., Kato N., Ohkoshi S., Shibuya A., Shimotohno K.;
Nakazawa T., Kato N., Ohkoshi S., Shibuya A., Shimotohno K.;
-1. BUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
Interpro: IPRO02522; HCV_capsid.
Pfam: PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ohno T., Mizokami M.; *Determination of nine genotypes of hepatitis C virus using PCR
                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
   Length 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 106;
                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
(Isolate USA8) genomic RNA (Genome polyprotein) (Fragment).
                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106 106
106 AA; 12001 MW; 25D0D5414B3EA9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         D945CB60661797E3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Score 154; DB 12;
; Pred. No. 9.9e-16;
0; Mismatches 0;
100.0%; Score 154; DB 12; 100.0%; Pred. No. 9.8e-16;
                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  125 AA
                                                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D49465; BAA08439.1; -.
InterPro; IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      125 AA; 14220 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-OCT-2002 (TrEMBLrel. 22,
                                                          28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 100 les 28; Conservative
                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-40271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein.
NON_TER 1
SEQUENCE 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-USA8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein.
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                      081831;
                                                                                                                                                                                                                                                                                                                           081831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9PXN3
                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TJUSL 3
                                                                                                                                                                                                                                                              RESULT 081831 1D 088 AC 088 01 DT 01 DT 01 DT 01 DT 01 DE (15 DE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

ô

Gaps

ö

Length 191;

```
STRAIN-JT';

MEDLINE-92295714; PubMed-1318627;

Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T.,

Hijikata M., Ishimura Y., Shimotohno K.;

"Molecular cloning of hepatitis C virus genome from a single Japanese
"ancier: sequence and an antiation within the same individual and among
infected individuals.";
single Japanese carrier in Nagasaki prefecture and genome analysis of El and E2/NS1 envelope glycoprotein regions.";
Jpn. J. Trop. Med. Hyg. 22:169-177(1994).

-1. SUBUNIT: THE VIRION OF THIS VIRIUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D63857; BAA09919.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane SEQUENCE 3010 AA; 327503 MW; C7BDB38169D6B3CF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein.
Heparitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    081541;
01-NOV-1996 (T.FEMBLrel. 01, Created)
01-NOV-1996 (T.FEMBLrel. 01, Last sequence update)
01-MAR-2003 (T.FEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 154; DB 12; 100.0%; Pred. No. 3.7e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; 1FR001095; RNA_pol_DS_FS.
InterPro; IPR007095; RNA_pol_DS_FS.
InterPro; IPR007095; RNA_pol_DS_FS.
InterPro; IPR007094; RNA_pol_DS_FS.
InterPro; IPR007094; RNA_pol_DS_FS.
InterPro; IPR007094; RNA_pol_DS_FS.
Pfam; PF01560; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01006; HCV_NS1; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01001; HCV_NS53; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01006; HCV_NS54; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PSS0507; RDRP_POSITIVE; 1. PROSITE; PSS0501; RDRP_VIRAL; 1.
                                                                                                                                                                                                InterPro; IPRO0252; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_MOV.
InterPro; IPR002531; HCV_MS2.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR00149; HCV_MS4.
InterPro; IPR00149; HCV_MS4.
InterPro; IPR002868; HCV_MS5.
InterPro; IPR002868; HCV_MS5.
InterPro; IPR002868; HCV_MS5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Pre-
                                                                                                                                                                                   InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 081541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                            PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
COST FOOTEIN; PS50521; RDRP_VIRAL; 1.
COST PROTEIN; Envelope Protein; Glycoprotein; Nonstructural protein;
PPOLYPROTEIN; RNA-directed RNA polymerase; Transmembrane,
SEQUENCE 3010 AA; 326793 MW; 3D89304314F9F795 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zheng W.Z.; "Genotype identification of hepatitis c virus (HCV) isolated from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel: 01, Created)
01-NOV-1996 (TrEMBLrel: 01, Last sequence update)
01-NAR-2003 (TrEMBLrel: 23, Last annotation update)
El and E2/NS1 envelope glycoprotein (Genome polyprotein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zheng W.-Y.Z.; Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 154; DB 12; 100.0%; Pred. No. 3.7e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                      InterPro; IPROUA45; NOV.NS4a.
InterPro; IPRO01499; HCV_NS4a.
InterPro; IPR001869; HCV_NS5a.
InterPro; IPR001869; HCV_NS5a.
InterPro; IPR001869; HCV_NS5a.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007091; HCV_CORP; I.
Pfam; PF01560; HCV_NS1; I.
Pfam; PF01006; HCV_NS3; I.
Pfam; PF01006; HCV_NS3; I.
Pfam; PF01006; HCV_NS4a; I.
Pfam; PF01006; HCV_NS4a; I.
Pfam; PF0156; HCV_NS5a; I.
Pfam; PF0156; HCV_NS5a; I.
Pfam; PF01006; HCV_NS4a; I.
Pfam; PF0100998; Viral_RAP; I.
Pr0Dom; PD186662; HCV_NS1; I.
                                                                                                                                                                                                    Interpro; IPR001410; DEAD.
Interpro; IPR002522; HCV_capsid.
Interpro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                          InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 100.
nes 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
      SEQUENCE FROM N.A STRAIN-HCV-S1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-HCV-N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                             Lim S.P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q81989
Q81989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
Q81989
```

SO ON WENT OF STREET OF ST

ô

6

ô

Gaps

ö

```
Songsivilal S., Kanistanon D., Kunkitti R., and isolates of hepatitis {\tt C} "identification and characterisation of Thai isolates of hepatitis {\tt C}
Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.

-I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN B. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, D49459; BAA21059.1; -.

InterPro; IPR002522; HCV_capsid.

Pfam, PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     virus.;
Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-- SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLECCARSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
-- PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF EMBL, U23745; AA455052.1;
-- InterPro; IPR002222; HCV.capsid.
-- Profess of the Complex of the Complex of Embl. (1974); AA455052.1;
-- Profess of the Complex of the Complex of Embl. (1974); AA65052.1;
-- Profess of the Complex of the Complex of the Complex of Embl. (1974); AA65052.1;
-- Profess of the Complex of the 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                         Length 109;
                                                                                                                                                                                                                                                                                                          109
12318 MW; 9999DBD249BF93BO CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Last Sequence update) 01-COT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment). Heparlis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                      Score 150; DB 12;
Pred. No. 4.1e-15;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 146; DB 12;
Pred. No. 6.5e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    45 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRQTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                      97.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45 45
45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94.8%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 96.49
Watches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                    109 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-HCV-BB21;
                                                                                                                                                                                                                                                                                                          109
                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER 10
SEQUENCE 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q68307
Q68307;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          968308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          068308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      068307
                          ZCCCCC ZE ZEZ ZE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NX OCC DE PETE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
      Virus Res. 23:39-53(1992).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGNS.
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, D11355; BAA18894.1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         v polymerase; Transferase; Transmembrane. PUT. P22. PUT. GP35. PUT. GP70. PUT. NS2. PUT. NS3. PUT. NS3. PUT. NS3. PUT. NS3. PUT. NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane
CHAIN 1 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NoV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
(isolate SLI) genomic RNA (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ohno T., Mizokami M.; "Determination of nine genotypes of hepatitis C virus using PCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         05F0B2102CF9DD9D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 th 100.0%; Score 154; DB 12; SImilarity 100.0%; Pred. No. 3.7e-14; 28; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS4A.
NS4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR00149; HCV_NS4a.
InterPro; IPR01409; HCV_NS4b.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR002166; HCV_RdRP.
InterPro; IPR007099; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_caps1d InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         326564 MW;
                                                                                                                                                                                                                                                                                                                             InterPro; IPR002519; HCV_env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR004109; HCV_NS3
                                                                                                                                                                                                                              InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2014 301
3010 AA;
                                                                                                                                                                                                       P26663; 1JXP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-SL1;
                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mehtod."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  081807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
081807
081807
AC 08188
DD 01-N0
DD 01-N
```

a

ö

Gaps

ó

g

```
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                     Polyprotein.
NON_TER 60 60
SEQUENCE 60 AA; 6725 MW;
                                                                                                                                                                                                                                                      94.88;
96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6881 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94.8%;
96.4%;
                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 96.4'
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
61 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 27; Conserv
                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                   NCBI_TaxID-11103;
 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-RIG256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-2002
01-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QBJYR4
QBJYR4;
                                                                                                                                                                                                                                                                                                                                                                                      Q8JYS3
                                                                                                                                                                                                                                                                                                                                                               RESULT 13
Q8JYS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8JYR4
ID Q8
AC Q8
g
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                             Gaps
                                                                      ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Songsivilal S., Kanistanon D., Kunkitti R., Italiacation and characterisation of Thai isolates of hepatitis {\tt C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
       STRAIN-HCV-BB38;
Songsivilai S., Kanistanon D., Kunkitti R.;
"Identification and characterisation of Thai isolates of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: U23748; AAA65055.1;
                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                     Length 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 45;
                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                             CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CC527167096DDAF6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 146; DB 12;
Pred. No. 6.5e-15;
0; Mismatches 1;
                                                                                                                                                                                                    Score 146; DB 12;
Pred. No. 6.5e-15;
                                                                                                                                                                                                                                                                                                                                      45 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60 AA.
                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                FILLIFILITI FILLIFILIFILITI S PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522; HCV_capsid. Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-0CT-2002 (TrEMBLrel. 22, Cr
01-0CT-2002 (TrEMBLrel. 22, La
01-MAR-2003 (TrEMBLrel. 23, La
Genome polyprotein (Fragment).
                                                                                                                                                                                                    94.8%;
96.4%;
                                                                                                                                                                            45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45 45
45 AA; 5014 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best_Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                         Query Match
Best Local Similarity 96.4°
                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-HCV-C94009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                     068310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08JYS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08JYS2
                                                                                                                                                                                                                                                                                                             RESULT 11
Q68310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
QBJYS2
ID QBJYS
AC QBJYS
DT 01-OC
DT 01-OC
DT 01-MAC
```

STRDR

g ç

```
Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
Submitted (APR-2002) to the EMBL/Genbenk/DDBJ databases.
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIDOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL; AF506612; AAM3388.1; -..
InterPro, IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                           BY A
                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                           9
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                 Length 60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 12; Length 61;
                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                            ACCE7D9C8B90299A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B92CCE7D9C8B9029 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-OCT-2002 (TrEMBLrel. 22, Last sequence update) 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                             Score 146; DB 12;
Pred. No. 8.9e-15;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 146; DB 1;
Pred. No. 9e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 9e-1
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORE PROTEIN.
                                                                                                                                                                                                                                                                                                                             1 PKPORKTKRNTYRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                  73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 22, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
```

ö

us-09-491-146a-25.rspt

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ralinina O., Norder H., Mukomolov S., Magnius L.O.;

*A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg...;

J. Virol. 76:4034-4043(2002).

-1. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: AV701080; AALS854.1;

InterPro; IPR002522; HCV_capsid.

Profile AND AND ALS854.1;

InterPro; IPR002522; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                            Shustov A.V., Gavrilova I.V., Netesov S.V.;
"Genetic variability of hepatitis C virus in Western Siberia.";
submitted (Apr.202) to the EMBL/Geneank/DBBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AF506621; AAAM33397.1; -
Interpro; IPR002522; HCV_capsid.
Pfan. PF01543; HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
01-CCT-2002 (TREMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 94.8%; Score 146; DB 12; Length 7 Best Local Similarity 96.4%; Pred. No. 1.1e-14; Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 AA; 11236 MW; 12AD90F5F5885AE6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6EDC082DAE8CB796 CRC64;
01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 146; DB 12;
Pred. No. 1.5e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-21904745; Pubmed-11907242;
                                                                                               Genome polyprotein (Fragment).
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73 73
73 AA; 8201 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 96.4'
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. STRAIN-RIG291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                        Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein.
NON_TER
NON_TER
SEQUENCE 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                080P85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
080298
10 08029
AC 08029
DT 01-JU
DT 01-OC
DE COTE
OC VITUS
CO HEDAC
OX HE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sor
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

August 7, 2003, 11:18:58 Search completed: Augus Job time: 26.6364 secs

```
0, App
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 25, App.
                                                                                                                                                                                                          (without alignments)
112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                            August 7, 2003, 11:07:41; Search time 10.5455 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            328717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
/cgn2_6/ptodata/1/laa/5B_COMB.pep:*
/cgn2_6/ptodata/1/laa/6A_COMB.pep:*
/cgn2_6/ptodata/1/laa/6B_COMB.pep:*
/cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-635-886C-192
US-08-291-887-23
US-08-290-665A-187
US-08-290-665A-189
US-08-290-665A-190
PCT-US95-10398-187
PCT-US95-10398-187
PCT-US95-10398-187
US-08-380-160-2
US-08-380-160-2
US-08-380-160-2
US-08-380-160-2
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -US94-07088-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-836-075A-10
                                                                                                                                                                                                                                                                                                                    154
1 PKPQRKTKRNTYRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1-US93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                     protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                           US-09-491-146A-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match Length DB
                                                                                                                                                                                                                                                                                           Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sed
                                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB s
Maximum DB s
                                                                                                                        protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                                      Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ĕ
```

```
Sequence 25, Application US/08921887

Patent No. 6030771

GENERAL INFORMATION:
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
TITLE OF INVENTION: BNDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAN
CORRESPONDENCE ADDRESS 5

CORRESPONDENCE & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY. Atlanta
            Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 60, Appli
Sequence 1, Appli
Sequence 15, Appli
                                                                                                                                                                                                                                                                Sequence 157,
Sequence 158,
Sequence 159,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 154; DB 3;
100.0%; Pred. No. 1.9e-15;
iive 0; Mismatches 0;
                                               US-08-904-686A-8
US-08-315-850-8
US-08-501-195-2
US-08-836-075A-46
US-08-836-075A-60
US-08-501-195-1
US-08-51-135-2
US-08-95-135-2
US-08-07-681-701-16
US-08-07-681-701-16
US-08-07-681-701-16
US-08-08-290-655A-155
US-08-290-665A-155
                                                                                                                                                                                                                                                             US-08-290-665A-157
US-08-290-665A-158
US-08-290-665A-159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFRX: 404-818-3799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS: not relevant TOPOLOGY: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
STRANDEDNESS: no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
ZIP: 30303-1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 28; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ANTI-SENSE: NO
FRAGMENT TYPE: 1
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; ORGANISM:
US-08-921-887-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
```

Gaps

ö

Indels

Conservative

å g

```
LIGATION METHOD FOR MAKING THE SAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R.H.
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 28;
                                                                                                                                                                                                                                                                                                                                  COUNTY: USA
ZIP: 30303-1769
ZORPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMpatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,887
                                                                                                                                                                                                                                               ADDRESSEE: JONES & ASKEW, LLP
STREET: 191 Peachtree Street, N.W., 37th Floor
CITY: Atlanta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              94.8%; Score 146; DB 3; Lv 96.4%; Pred. No. 2.7e-14; Live 0; Mismatches 1;
                                                                                                       APPLICANT: KHUDYAKOV, YURI E.
APPLICANT: FIELDS, HOWARD A.
TITLE OF INVENTION: MOSAIC PROTEIN AND RES
TITLE OF INVENTION: ENDONUCLEASE ASSISTED
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: BURH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: AMINO ACID SEQUENCES
TITLE OF INVENTION: CORE GENES OF ISOLAT
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
TITLE OF INVENTION: SEQUENCES IN DIAGNOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKTKRNTYRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03063-0380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 187, Application US/08290665A Patent No. 5882852
                                            Sequence 23, Application US/08921887
Patent No. 6030771
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 0306:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: WARREN, WILLIAM L.
REGISTRATION NUMBER: 36,71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; ORGANISM: Hepatitis virus US-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 28 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ANTI-SENSE: NO ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-290-665A-187
                         JS-08-921-887-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  đ
                                                                                                                                                                                       GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geart
APPLICANT: DELESS, Robert
APPLICANT: DELESS, Robert
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
TITLE OF INVENTION: VIRUS
FILE REPERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT PAPLICATION NUMBER: DET/EP94/03555
PRIOR FILING DATE: 1996-10-28
PRIOR FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PATCHING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SEQ ID NO 191
LENGTH: 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: DELEXS, Robert
APPLICANT: DELEXS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNDOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: UNRUS
FILE REFERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PALENTIN VERSION 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 154; DB 4;
100.0%; Pred. No. 3.3e-14;
11ve 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%; Score 154; DB 4;
100.0%; Pred. No. 3.3e-14;
Live 0; Mismatches 0;
1 PKPORKTKRNTYRRPODVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 32
                       Sequence 192, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
APPLICANT: LEROUX-ROBELS, Geert
APPLICANT: DELEYS, Robert
                                                                                                                                                    Sequence 191, Application US/08635886C Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSGANISM: hepatitis C virus US-08-635-886C-192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 100.C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best_Local Similarity 100.
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
US-08-635-886C-192
                                                                                                                                 US-08-635-886C-191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NO 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
```

å g

g ŝ

ö

Gaps

```
APPLICANT: PURCELL, R.H.

TITLE OF INVENTION: UNCLECOTIDE AND DEDUCED
TITLE OF INVENTION: AUCLECOTIDE SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
NUMBER OF SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                 DB 2; Length 191; 1.9e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                              Query Match 94.8%; Score 146; Best Local Similarity 96.4%; Pred. No. 1 Matches 27; Conservative 0; Mismatch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94.8%; Score 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION UNBER: US/08/290,665A
FILING DATE: 15-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/CDCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: BUKH, J., MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 190, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               190:
                                                                                           188:
  TELECOMMUNICATION INFORMATION:
                                                            INFORMATION FOR SEC ID NO: 186
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK12
US-08-290-665A-190
                          (212) 758-48C
(212) 751-6849
                                                                                                                                                                                                                                                     ORGANISM: homosapiens;
INDIVIDUAL ISOLATE: S52
US-08-290-665A-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX: 421792
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: un
ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-290-665A-190
                          TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                           TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 188, Application US/08290665A

Patent No. 5882852

GENERAL INFORMATION:
APPLICANT: BUCH, J., MILLER, R.H. AND
APPLICANT: BUCKELL, R.H.
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEBRITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: ASQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 94.8%; Score 146; DB 2; Lu
Best Local Similarity 96.4%; Pred. No. 1.9e-13;
Matches 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                         CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REPERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQKKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                          COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,665A
FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/290,665A FILING DATE: 15-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HK10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
                                                                                           MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
                                                                   COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: NEW YORK
STATE: NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
US-08-290-665A-188
                                           10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-290-665A-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM:
STATE: N
COUNTRY:
```

ò a

```
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND CORE GENES OF ISOLATES OF HEPATITIS C VIRUS AND THE USE OF REAGENTS DERIVED FROM THESE 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Match 94.8%; Score 146; DB 5; Local Similarity 96.4%; Pred. No. 1.9e-13; les 27; Conservative 0; Mismatches 1
                                                                     NUCLEOTIDE AND DEDUCED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                               MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 190, Application PC/TUS9510398
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                       ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: RICHARD W. BORK
REGIESTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 201
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                      IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEX: 421792
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: WORDPERFECT 5.1 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (212) 759-48(
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ) ORGANISM: homosapiens
) INDIVIDUAL ISOLATE: S52
PCT-US95-10398-188
                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                          ZIP: 10154
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
                                                                   TITLE OF INVENTION: NO TITLE OF INVENTION: AN TITLE OF INVENTION: CO TITLE OF INVENTION: AN TITLE OF INVENTION: SEQUENCES: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC OPERATING SYSTEM:
                                                                                                                                                                                                                                                             NEW YORK
NEW YORK
        GENERAL INFORMATION:
APPLICANT: BUKH,
                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps .
                          Gaps
                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED

TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94.8%; Score 146; DB 5; Length 191;
96.4%; Pred. No. 1.9e-13;
tive 0; Mismatches 1; Indels
  Pred. No. 1.9e-13;
0; Mismatches 1; Indels
                                                                 1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                     MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIELE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-ANG-1995
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RIGIARD W: BORK
REGISTRATION NUMBER: 36,428
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RIGIARD W: BORK
REGISTRATION NUMBER: 36,429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: BUKH, J., MILLER, R.H. AND APPLICANT: PURCELL, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
PCT-US95-10398-188
; Sequence 188, Application PC/TUS9510398
                                                                                                                                                                                                                Sequence 187, Application PC/TUS9510398 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                SSEE: MORGAN & FINNEGAN P.: 345 PARK AVENUE NEW YORK NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEX: 421792
INFORMATION FOR SEQ ID NO: 187: SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: (212) 758-4800
TELEFAX: (212) 751-6849
  96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: HK10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
Best Local Similarity 96.49
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                            -US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-10398-187
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
STREET: 34
                                                                                                                                                                                                                                                                                 APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                        RESULT
                                                                                                       a
                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: LERONY-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: DELEYS, Robert
TITLE OF INVENTION: URMUNDCMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
TITLE OF INVENTION: VIRUS
FILE REFERENCE: 2752-18
CURRENT FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: DCT/EP94/03555
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR PILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: PatentIn version 3.1
SEQ ID NO 230
LENGTH: 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , DB 3;
3.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Score 146; DB
; Pred. No. 3.2e
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (144)..(144)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION: (156)..(157)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             any amino acid
                                                                  FILING DATE: 23 Oct 1955

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: 1NNS:004
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
FILING DATE: 21 Apr 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Sequence 230, Application US/08635886C
; Patent No. 6555114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 94.8%;
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 319 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION: (161)..(161)
OTHER INFORMATION: Xaa 1s
                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MISC_FEATURE LOCATION: (156)..(157)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE LOCATION: (167)..(167)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 44, Application US/08836075A
Patent No. 6180768
GENERAL INFORMATION:
APPLICANT: MAERTENS, GERT
APPLICANT: STUTYER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match

94.8%; Score 146; DB 5; Length 191;
Best Local Similarity 96.4%; Pred. No. 1.9e-13;
Matches 27; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IN PC COMPATIBLE
SYSTEM: PC-DOS/MS-DOS
Microsoft Word 6.0 / ASCII text output
                                                                                                                                                                                                                                                         COURENT APPLICATION DATE:

CARENT PEPLLCATION DATE:

FILING DATE: 15-AUG-1995

CLASSIFCATION

PROPERTY APPLICATION DATE:

APPLICATION NUMBER: 08/086,428

FILING DATE: 29 JUNE 1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/290/665

FILING DATE: 15 AUGUST 1994

ATTONNEY, AGGNT INFORMATION:

NAME: RICHARD W. BORK

REGISTRATION NUMBER: 36,459

REFERENCE/DOCKT NUMBER: 36,459

REFERENCE/DOCKT NUMBER: 36,459

REFERENCE/DOCKT NUMBER: 36,459

REFERENCE/DOCKT NUMBER: 36,459

FELECHONE: (212) 759-4800

TELECHONE: (212) 759-4800

TELECHONE: (212) 751-6849

TELECHONE: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,075A
                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: DK12
PCT-US95-10398-190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Floppy disk
  STREET: 345 PARK AVENUE CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: unknown ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM:
SOFTWARE: Microso
                                                     NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CITY: HOUSTON STATE: TEXAS
                                                                            USA
                                                                                                10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-836-075A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER:
```

à Q

Gaps

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                    Score 145; DB 3; Length 34;
Pred. No. 4.5e-14;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-024045
FILING DATE: 09-FEB-1996
ATTORNEX/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 145; DB 4; L
Pred. No. 5.7e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: YAMAGUCHI, Kenjiro
APPLICANT: KASHIMAKUMA, Tomiko
APPLICANT: CHIBA, Yukie
APPLICANT: CHIBA, Yukie
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAMA, AKIRa
TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                           1 PKPORKTKRNTYRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                    NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 053466/0225
TELECOMMUNICATION:
                                                                                                    ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 36, Application US/09020846
; Patent No. 6322965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3000 K Street, N.W
                                                                                                                                                                                                                               94.2%;
illarity 96.4%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              94.2%;
                       sss: single
linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        43 amino acids
                                              TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
            TYPE: amino acid
STRANDEDNESS: sir
                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: DALBON, PASCAL
APPLICANT: JOLIVET, MIChAL
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
TITLE OF INVENTION: FOR DETECTING THE LATTER
CORRESPONDENCE: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                             94.8%; Score 146; DB 4; Length 319; 96.4%; Pred. No. 3.2e-13; 1ve 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTIRRPQDVKFPGGGQIVG 32
                                                                                            FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (174)..(174)
OTHER INFORMATION: Xaa is any amino acid
INFORMATION: Xaa is any amino acid
                                                    LOCATION: (171)..(172)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                            NAME/KEY: MISC_FEATURE LOCATION: (177)...(177) OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                              LOCATION: (232)..(232)
OTHER INFORMATION: Xaa is any amino acid
                                                                                                                                                                                                                                                                                                                                                                            LOCATION: (233). (233)
OTHER INFORMATION: Xaa is any amino acid
08-635-886C-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: WPB 28682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Sequence 6, Application US/08380160
; Patent No. 6235284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 96.4*
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                       NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                      FEATURE:
NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: MISC_FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
Search completed: August 7, 2003, 11:23:50
Job time : 10.6364 secs
                                                                                                                                                                                                                          Michel
SYNTHETIC POLYPEPTIDES BELONGING TO THE
SERPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY
FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94.2%; Score 145; DB 3; Length 44; 96.4%; Pred. No. 5.9e-14; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:

WEDDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
FILING DATE:
CLASSIFICATION S30
PRIOR PAPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTONNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 30,024
REGISTRATION NUMBER: 30,024
REGISTRATION NUMBER: 30,024
REGISTRATION NUMBER: 30,024
RELEPHONE: (703)836-6400
TELEFAK: (703)836-2787
1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                     Sequence 1, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
TITLE OF INVENTION: SYNTHERIC POLY
TITLE OF INVENTION: HEPATITIS C VI
TITLE OF INVENTION: POR DETECTING
NUMBER OF SEQUENCES: 12
                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. Box 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 96.43
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCATION: 1..44

COTHER INFORMATION: COTHER INFORMATION: COTHER INFORMATION: COS-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                ۲
                                                                                                                                                                                                                                                                                                                                                                                                            STATE: VA
  ò
                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

g

Sequence Sequence Sequence

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Sequence 1, App Sequence 12, App Sequence 12, App Sequence 12, App Sequence 50, App Sequence 51, App Sequence 52, App Sequence 52, App Sequence 54, App Sequence 50, App Sequence 21, App Sequence 21, App Sequence 22, App Sequence 4, Appl Sequence 4, Appl Sequence 1, Appl Sequence 2, Appl Sequence 2, Appl Sequence 3, Appl Sequence 3, Appl Sequence 2, Appl Sequence 3, Appl Sequence 3, Appl Sequence 2, Appl

## ALIGNMENTS

US-09-952-572-9 US-09-929-955-1 US-09-747-19-20 US-09-891-894-3 US-10-259-25-20 US-10-259-25-20 US-09-238-076-2

3011 3011 3012

TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGN ADDRESSER: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210-443
COMPUTER READALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATA:
PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
PELLORION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995 NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775 REFERENCE/DOCKET NUMBER: INNS:004 INFORMATION FOR SEQ ID NO: 44: // Sequence 44, Application US/09851138
// Publication No. US20020183508A1
// GENETEL INFORMATION:
APPLICANT: MAERTENS, GEERT
// APPLICANT: MAERTENS AGENTS ATTORNEY/AGENT INFORMATION CORRESPONDENCE ADDRESS NUMBER OF SEQUENCES: US-09-851-138-44 RESULT 1

Cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/PCTUG\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.ppp:\*

(cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.ppp:\* Published\_Applications\_AA:\* Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Database :

length: 0 length: 2000000000

sed

Minimum DB Maximum DB

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```
APPLICANT: HOWARD ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C \
TITLE OF INVENTION: PROFIEINS
FITLEOF INVENTION: PROFIEINS
FILE REFERENCE: 14114.034902
CURRENT APPLICATION NUMBER: US/09/758,308
CURRENT APPLICATION NUMBER: 60/092,339
PRIOR APPLICATION NUMBER: 60/092,339
PRIOR RILING DATE: 1999-07-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PATENTIN VERSION 3.0
SEC ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                 COMPUTE: IBM PC Compatible
COMPUTE: IBM PC Compatible
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94.2%; Score 145; DB 10;
96.4%; Pred. No. 1.7e-13;
tive 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 145; DB 9;
Pred. No. 2.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775
                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1, Application US/09758308 Patent No. US20020090607A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGIH: 74 amino acids
                                                                  STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO: 10
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                ADDRESSEE: ARNOLD,
                       CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: MAERTENS, GEERT STUYVER, LIEVEN STUYVER, LIEVEN TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: DELIAN, FRANCOIS
APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: LAURENT
APPLICANT: LAURENT
APPLICANT: LACONX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
TITLE OF INVENTION: TREATING AN HCV INFECTION
FILE REFERENCE: 103959
CURRENT APPLICATION NUMBER: US/09/389,756
PRIOR APPLICATION NUMBER: US/09/389,756
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PRIOR FILING DATE: 1999-09-07.
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 1
LENGTH: 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                            Length 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 44;
                                                                                                                                                                                                                      Indels
                                                                                                                                                                        Query Match 94.8%; Score 146; DB 10; Best Local Similarity 96.4%; Pred. No. 6.2e-13; Matches 27; Conservative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 145; DB 15;
Pred. No. 9.9e-14;
0; Mismatches 1;
  LENGTH: 319 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 44:
                                                                                                                                                                                                                                                               1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                      4 PKPQRKIKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 10, Application US/09851138 Publication No. US20020183508A1
                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1, Application US/10367677 iblication No. US20030118604A1 ENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGENTS
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94.2%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: JOLIVET, MICHEL APPLICANT: PENIN, FRANCOIS APPLICANT: DALBON, PASCAL APPLICANT: LADAVIERE, LANE APPLICANT: LACOUX, XAVIER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAGES: 3392-3396
DATE: 1991
; RELEVANT RESIDUES: 2 TO 45
US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 96.4'
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
US-10-367-677-1
                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

Gaps

ö

```
APPLICANT: MAERTENS, GEERT
APPLICANT: MAERTENS, LIEVEN
STUYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: TEXAS

CONTRY: USA
2 IP: 77210-443

ZONDUTRY: USA
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION UNDER: US/09/851,138
FILING DATE: 09-MMY-2001
PRIOR APPLICATION DATA:
                                                                      Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 108;
                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                               Score 145; DB 10;
Pred. No. 2.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.2%; Score 145; DB 10; 96.4%; Pred. No. 2.6e-13;
                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                            1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 14, Application US/09851138; Publication No. US20020183508A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 78, Application US/09921397
Patent No. US20020151484A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHIT
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 108 amino acids TYPE: amino acid
                                                               94.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
; ORGANISM: Hepatitis C virus US-09-921-397-77
                                                                                                          27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 96.49 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO
                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                            RESULT 7
US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                          Matches
                                                                                                                                                      ö
                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 77, Application US/09921397

Patent No. US2020151484A1

GENERAL INFORMATION:

TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a TITLE OF INVENTION: applications thereof

TITLE OF INVENTION: applications thereof

TITLE OF INVENTION: applications thereof

CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT APPLICATION NUMBER: EP 00402225.7

PRIOR APPLICATION NUMBER: EP 0040225.7

PRIOR APPLICATION NUMBER: 2001-08-03

WUMBER OF SEQ ID NOS: 156

SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 97;
                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM: OCCOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 145; DB 10;
Pred. No. 2.3e-13;
0; Mismatches 1;
                                                                                                                                      APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
CITY: Mashington
STATE: D. C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,721
FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB93/00410
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/COCKET NUMBER: 1808-157A
TELECOMMUNICATION:
TELECOMMUNICATION:
TELEPHONE: (202)/83-6040
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/09/756,875
                                                                           Sequence 8, Application US/09756875
Patent No. US20020150990A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-756-875-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                   D. C.
                                                                                                                                                                                                                                                                                                                     COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-921-397-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
```

ö

Gaps

Page

ö

```
STUTVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                   ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 145; DB 10;
Pred. No. 3.4e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: 08/836,075
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
NAME: KAMBRER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                         1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 152, Application US/09899046
; Publication No. US20030008274Al
; GENERAL INFORMATION:
                                                                                                                                                                                                                    Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 138 amino acids
                ö
                                                                                                                                                                                                                                                                                       APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 20'
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 96.4
Matches 27; Conservative
            27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
US-09-899-046-152
                                                                                                                                                                      RESULT 10
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-851-138-60
              Matches
                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AFFLICANT: FASTILE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                            ö
TITLE OF INVENTION: SID nucleic acids and polypeptides selected TITLE OF INVENTION: pathogenic strain of the hepatitis C virus FITLE OF INVENTION: applications thereof FILE REPERENCE: B4809A - JA2 CURRENT APPLICATION NUMBER: US/09/921,397 CURRENT APPLICATION NUMBER: US/09/921,397 CURRENT FILING DATE: 2001-08-02 PRION FILING DATE: 2000-08-03 NUMBER OF SEQ ID NOS: 156 SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPTE: 16AA3

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
APPLICATION DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: EP 9487016.9
FILING DATE: 20 Oct 1994
APPLICATION NUMBER: EP 9487016.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORREY/AGENT INFORMATION:
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORREY/AGENT INFORMATION:
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORREY/AGENT INFORMATION:
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 JUN 1995
ATTORREY/AGENT INFORMATION:
APPLICATION NUMBER: EP 95870076.7
REGISTRATION NUMBER: 29,775
REGISTRATION NUMBER: 29,775
                                                                                                                                                                                                                                                                                                                                                                                 Length 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                            1; Indels
                                                                                                                                                                                                                                                                                                                                                                         Score 145; DB 10;
Pred. No. 2.8e-13;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 145; DB 10;
Pred. No. 3.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-851-138-46
; Sequence 46, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 137 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94.28;
                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                   ; ORGANISM: Hepatitis C virus US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                             LENGTH: 113
                                                                                                                                                                                                                                     SEQ ID NO 78
                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ì
```

Gaps

ö

Indels

```
New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy. 270
                                                                                                                                                                                                                                                                                                                                                                                             ; DB 11; Length 169; 4.3e-13;
      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READBLE FORM:
MEDIUM TYPE: FIOPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94.2%; Score 145; DB 11; 96.4%; Pred. No. 4.3e-13;
                                                                                                                                                                                                                                                                                                                                                                                             Score 145; DE
Pred. No. 4.36
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 44, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
US-09-878-281-42
Squence 42, Application US/09878281
Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                        TYPE: amino acids TOPOLOGY: ):-
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                         ) MOLECULE TYPE: protein US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  සි
                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
NUMBER OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: New sequences of hepatitis C virus TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy. NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                 Score 145; DB 11; Length 166;
Pred. No. 4.2e-13;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
94.2%; Score 145; DB 11;
Best Local Similarity 96.4%; Pred. No. 4.2e-13;
Matches 27; Conservative 0; Mismatches 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PKPQRKTKRNTYRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/09/878,281 FILING DATE: PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 42, Application US/09899046 Publication No. US20030008274A1 GENERAL INFORMATION: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acids
                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLGGY: linear
                                                                                                                                                                                                                                                                                                                                                                   Query Match

Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-878-281-152
                                                                                                                                                                                                                                                                                                   ) MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

ö

Gaps

ö

Indels

```
us-09-491-146a-25.rapb
```

Search completed: August 7, 2003, 12:01:12 Job time: 14.3636 secs

us-09-491-146a-26.rpr

protein search, using sw model

protein

Run on:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

US-09-491-146A-26

Perfect score:

Sequence:

```
Suppose polyprotein - hepatitis C virus (genotype 2, N2) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety; genotype 2, N2
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: 34153
R;Van Doorn, L.J.; Rater, G.E.M.; Brouwer, J.T.
Submitted to the EMBL Data Library, January 1994
A;Bescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: S41341
A;Accession: S41353
A;Molecule type: genomic RNA
A;Residues: 1-108 cVAN>
A;Cross.references: EMBL:229456
A;Experimental source: genotype 2, N2
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein #status predicted <NAT>
                                             polyprotein
   otide - hepa
polyprotein
                                   polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nicontains: core protein
C; Species: hepatitis C virus
A;Variety: genotype 2, N4
A;Variety: genotype 2, N4
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41355
R; Van Doorn; L.J.: Kleter, G.E.M.: Brouwer, J.T.
submitted to the RmBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                               genome
genome
genome
genome
genome
genome
genome
genome
                                                                                                                                                                                                         genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (genotype 2, N4) (fragment)
                   qenome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142; DB 2; Length 108;
Pred. No. 1.4e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 142; DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A, Molecule type: genomic RNA
A, Residues: 1-108 <VAN>
A,Cross.references: EMBL:229458
A,Experimental source: genotype 2, N4
C, Superfamily: hepatitis C virus genome polyprotein
C, Keywords: capsid protein; core protein; polyprozein
F,1-108/Product: core protein #status predicted <NAT>
                                                                                                                                                                                                                                                                                      ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PC2219
GNWVTC
GNWVCJ
S18030
                                                            GNWVCH
S40770
JC5620
JO1303
GNWVJ8
S41351
S41349
S41352
JO1584
GNWVC3
S21336
S41367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 92.9
Matches 26; Conservative
 876
3010
3010
3010
3010
3011
3033
115
115
115
640
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
991.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 2
S41355
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein - hepa
polyprotein - hepa
genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein
                                                                                                     August 7, 2003, 11:05:41; Search time 9.54545 Seconds (without alignments) 282.095 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
genome
genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               депоше
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                283308
             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                             tal number of hits satisfying chosen parameters:
                                                                                                                                                                                     156
1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                               283308 segs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
```

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_76:\* 1: pirl:\* 2: pir2:\* 3: pir3:\* 4: pir4:\*

Database :

seq length: 0 seq length: 2000000000

Minimum DB Maximum DB

ö

S12707 PC1284 JQ1925 JQ1926

JH0711 S19876

518032

\$41353 \$41355 \$41355 \$41356 \$41346 \$41341 \$4136 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340 \$41340

В

Length

Query Match 1

Result

Š

2

ö

```
RESULT
S41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                 Genome polyprotein - hepatitis C virus (genotype 2, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 2, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: $41357
Submitted to the EMBL Data Library, January 1994
A;Bescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: $41357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (genotype 1, N6) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N6
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: S41348
R;van Doorn, L.J.; Rleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
`Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
'eference number: S41341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (genotype 5, N5) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2; Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                         Indels
Pred. No. 1.4e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
A; Residues: 1-108 <VAN>
A; Cross-references: EMBL: 229451
A; Experimental source: genotype 1, N6
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-108/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   lecule type: genomic RNA
A;Cross-references: BMBL:229460
A;Experimental source: genotype 2, N6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; polyprotein
F;1-108/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match 91.0%; Score 142; DB 2; L
Best Local Similarity 92.9%; Pred. No. 1.4e-12;
Matches 26; Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       / Match 91.0%; Score 142; DB 2; I
Local Similarity 92.9%; Pred. No. 1.4e-12;
les 26; Conservative 0; Mismatches 2;
                                                                 1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKPORKPNRNTNRRPODVKFPGGGOIVG
95.98;
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Accession: S41348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                 ò
                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
genome polyprotein - hepatitis C virus (genotype 5, N4) (fragment)
N;Contains: core protein
C;Species: hepatitis, C virus
A;Variety: genotype 5, N4
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C;Accession: $4137,
Submitted to the EMBL Data Library, January 1994
A;Rescription: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A;Reference number: $41370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 1, N1) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
A;Variety: genotype 1, N1
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
A; Variety: genotype 5, N5
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: S41371
Submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41371
A; Accession: S41371
A; Residues: 1-112 <VAN>
A; Residues: 1-112 <VAN>
A; Cross-references: EMBL: Z29474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: $41341
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Bubmitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Accession: $41341
A; Accession: $41341
A; Accession: $41341
A; Accession: $1112 cVAN>
A; Residues: 1-112 cVAN>
A; Residues: 1-112 cVAN>
A; Cross-references: EMBL: 229444; NID: 9443850; PIDN: CAA82582.1; PID: 9443851
A; Cross-references: EMBL: 229444; NID: GCONDER CONDER COND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: genomic RNA
A;Residues: 1-114 <VAN>
A;Cross-references: EMBL:229473; NID:g443908; PIDN:CAA82611.1; PID:g443909
A;Experimental source: genotype 5, N4
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         atus p.c.,
Score 142; DB 2; Lenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142; DB 2; Le
Pred. No. 1.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Experimental source: genotype 5, N5
C; Superfamily: hepatitis C virus genome polyprotein
C; Neywords: capsid protein; core protein; polyprotein
F;1-112/Product: core protein *status predicted 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 91.0%; Score 142; DE
Best Local Similarity 92.9%; Pred. No. 1.5e
Matches 26; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
```

#text\_change 17-Nov-2000

LIPA.

ρλ

S

ð a

```
N; Contains: core protein
C; Species: hepatitis C virus
A; Variety: genotype 1, N2
C; Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
C; Accession: $4134
Submitted to the EMBL Data Library, January 1994
A; Reference number: $4134
A; Reference number: $41344
A; Reference number: $41344
                                                                                                                                                                                                                                                                                                                                                                 A) Cross references: EMBL:229445; NID:9443852; PIDN:CAA82583.1; PID:9443853 A) Experimental source: genotype 1, N10 C. Superfamily: hepetitis C virus genome polyprotein C; Superfamily: hepetitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F;1-115/Product: core protein #status predicted <NAT>
                        polyprotein - hepatitis C virus (genotype 1, N10) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein – hepatitis C virus (genotype 1, N2) (fragment)
                                                                                                                                                    C; Accession: S41342
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to
A; Reference number: S41341
A; Accession: S41342
                                              N.Contains: core protein
S.Species: hepatitis C virus
A.Variety: genotype 1, N10
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                  A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (genotype 5, N2) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
P,Variety: genotype 5, N2
Ate: 19-May.1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (genotype 5, N3) (fragment)
N;Contains: core protein
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Pariety: genotype 5, N3
C;Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
`.Accession: $41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Substituted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Reference number: S41341
A; Reference number: S41348
A; Rocession: S41368
A; Rocession: S41368
A; Residues: 1-14 < VAN.
A; Residues: 1-114 < VAN.
A; Residues: Bibl: Z29471; NID: g443904; PIDN: CAA82609.1; PID: g443905
A; Experimental source: genotype 5, N2
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein: Pill/Product: core protein #status predicted < WAIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           January John Dorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
Janitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.
A; Reference number: S41341
A; Reference number: S41341
A; Recession: S41369
A; Rocelule : 1-14 </P>
A; Residunes: 1-14 </P>
A; Residunes: BmBL: Z29472; NID: g443906; PIDN: CAA82610.1; PID: g443907
A; Experimental source: genotype 5, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: capsid protein; core protein; polyprotein
F;1-114/Product: core protein #status predicted <AAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 114;
                                                                          Length 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 114;
                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                          Score 142; DB 2; L. Pred. No. 1.5e-12; O; Mismatches 2;
C; Keywords: capsid protein; core protein; polyprotein F;1-114/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
91.0%; Score 142; DB 2;
Best Local Similarity 92.9%; Pred. No. 1.5e-12;
Matches 26; Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 142; DB 2;
Pred. No. 1.5e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%;
92.9%;
                                                                          ch 91.0%;
1 Similarity 92.9%;
26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 91.0
Best Local Similarity 92.9
Matches 26; Conservative
                                                                               Query Match
                                                                                                           Best Local
                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                       RESULT
S41369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

à

ö

Gaps

ö

Score 142; DB 2; Length 115; Pred. No. 1.5e-12; 0; Mismatches 2; Indels

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      #text_change 17-Nov-2000
A; Cross-references: EMBL: Z29447; NID: 9443856; PIDN: CAA82585.1; PID: 9443857
                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lipa.
                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (genotype 1, NB) (fragment) N:Contains: core protein N:Contains: core protein C. Species: hepatitis C virus A; Mariety; genotype 1, NB C:Date: 19-May-1994 #sequence_revision 26-Jul-1995 #text_change 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C; Accession: $41350
R; van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T.
submitted to the EMBL Data Library, January 1994
A; Description: Analysis of hepatitis C virus genotypes 1 to 5 by A; Reference number: $41341
A; Accession: $41350
                                                                                                                                                                    Length 115;
                                                                                                                                                                                                                     Indels
                            A; Experimental source: genotype 1, N2 (Sperfamily: hepatitis C virus genome polyprotein C; Keywords: capsid protein; core protein; polyprotein F; 1-115/Froduct: core protein *status predicted 
                                                                                                                                                              Score 142; DB 2; I
Pred. No. 1.5e-12;
); Mismatches 2;
                                                                                                                                                                                                                                                                                                            PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                     ö
                                                                                                                                                                    91.0%;
                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
```

10

RESULT

```
N.Contains: core protein
C.Species: hepatitis C virus
A.Variety: genotype 1, N5
C.Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Reference number: S41341
A; Reference number: S41341
A; Accession: S41347
A; Molecule type: genomic RNA
A; Residues: 1.115 < VAN>
A; Cross references: EMBL: 229450; NID: 9443862; PIDN: CAA82588.1; PID: 9443863
A; Experimental source: genotype 1, NS
A; Experimental source: genomy genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted < MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                         R)van Doorn, L.J.; Kleter, G.E.M.; Brouwer, J.T. submitted to the EMBL Data Library, January 1994
A.Description: Analysis of hepatitis C virus genotypes 1 to 5 by Lipa.
                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (genotype 1, N5) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 142; DB 2;
Pred, No. 1.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                        PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: August 7, 2003, 11:21:47 Job time : 9.54545 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: S41347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N'Contains: Core protein

Prodies: Parties C virus

riety: penotype 2, N3

c, ate: 19-May-1994 *sequence_revision 26-Jul-1996 *text_change 17-Nov-2000

C,Accession: $4134

Submitted to the EMBL Data Library, January 1994

A,Reference number: $4134

A)Reference number: $4134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      'ontains: core protein species and the species of t
A; Residues: 1-115 <VAN>
A; Cross-references: EMBL: Z29453; NID: 9443868; PIDN: CAA82591.1; PID: 9443869
A; Experimental source: genotype 1, N8
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:229448; NID:g443858; PIDN:CAA82586.1; PID:g443859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S41354
genome polyprotein - hepatitis C virus (genotype 2, N3) (fragment)
N:Contains: core protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nome polyprotein - hepatitis C virus (genotype 1, N3) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 115;
                                                                                                                                                                                                                                                                                Length 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 142; DB 2; Length ll. Pred. No. 1.5e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 142; DB 2; Length 11
Pred. No. 1.5e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                             2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Experimental source: genotype 2, N3
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: capsid protein; core protein; polyprotein
F; 1-115/Product: core protein *status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A:Experimental source: genotype 1, N3_
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; polyprotein
F:1-115/Product: core protein #status predicted 
                                                                                                                                                                                                                                                                            91.0%; Score 142; DB 2;
92.9%; Pred. No. 1.5e-12;
iive 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: EMBL: 229457
                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 92.9
les 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: genomic RNA
A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: genomic RNA A; Residues: 1-115 <VAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S41345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

Gaps

ö

1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28

ö

Gaps

ö

Length 115; Indels

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
```

OM protein - protein search, using sw model

August 7, 2003, 11:05:41; Search time 4.90909 Seconds (without alignments) 268.226 Million cell updates/sec Run on:

US-09-491-146A-26 156 1 PKPORKPURNTURRPQDVKFPGGGQIVG 28 Title: Perfect score:

Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 stal number of hits satisfying chosen parameters:

127863 seqs, 47026705 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0% Maximum Match 100%

SwissProt\_41:\* Database :

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		Description						P26663 h genome po	P26662 h genome po	ď	P26660 h genome po	P26661 h genome po	P26664 h genome po	c	ч	0	Q14738 h serine/th	P51816 homo sapien		028715 archaeoglob		P21573 xenopus lae	_	'n		P49415 drosophila	P97820 mus musculu	_	Q04572 shallot vir	P58405 rattus norv	P47708 mus musculu	P47709 rattus norv	Q9erg2 mus musculu	P25220 avian infec	Q9jmcl rattus norv
COLUMNIA		QI .	POLG_HCVJ2	POLG_HCVH4	POLG_HCVHK	POLG_HCVJ5	POLG_HCVJ7	POLG_HCVBK	POLG_HCVJA	POLG_HCVH	POLG_HCVJ6	POLG_HCVJ8	POLG_HCV1	POLG_HCVTW	POLG_HCVJT	2A5D_RABIT	2A5D_HUMAN	FMR2_HUMAN	SCD6_YEAST	YF57_ARCFU	GDF6_BOVIN	YB1_XENLA	RK4_TOBAC	CSP_PLABE	CSP_PLABA	SDC_DROME	M4K4_MOUSE	SC25_YEAST	COAT_SHVX	STR3_RAT	RP3A_MOUSE	RP3A_RAT	STR3_MOUSE	POLS_IBDVP	P5PA_RAT
		8	7	Н	٦	Н	ч	7	٦	٦	٦	٦	_	7	٦	П	٦	-	-	7	-	7	7	-	-	7	٦	-	-	٦	~	-	-4	-	Н
		Length	513	520	520	737	737	3010	3010	3011	3033	3033	3011	3010	3010	286	602	1311	349	167	436	303	282	339	347	395	1233	1252	262	461	681	684	196	993	1001
æ		Match	91.0	91.0	91.0	91.0	91.0	91.0			91.0							34.0	33.3	32.7	32.7	32.4	32.1	32.1	32.1	31.4	31.4		30.8					•	30.8
		Score	142	142	142	142	142	142	142	142	142	142	138	135	134	55.5	53.5	53	52	51	51	50.5	20	20	20	49	49	49	48	48	48	48	48	48	48
	Result	SO:	1	7	m	4	5	9	7	8	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

					P28276 herpes simp P10723 brugia mala	
POLS_IBDVS POLS_IBDVC	POLS_IBDVS C3G_DROME	P32_MYCGA XG_HUMAN	RNF4_HUMAN RS7_FUGRU	GRAB_MOUSE LIPA_CAUCR	IE63_HSV2H SYN_BRUMA	
				٦,		
1012	1012 1571	304 180	190	325	512 548	
30.8	30.8			30.1	30.1	
4 4 8 8	4 4 8 8	47.5	447	47	47	
44 IO	9 2	398	9:1	2.5	4 2	

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as along as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-922032; Pubmed-1314459;
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
A Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
Thomology to reported isolates: comparative study of four distinct genotypes.;
Thomology to reported isolates: comparative study of four distinct genotypes.;
Thomology 188:331-341(1992).

- 1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY FLAY A ROLE IN THE VIRAL RNA REPLICATION.
C -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                         01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66-noome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
Hepatitis C virus (isolate HC-32) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ΒX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              513 AA
                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_RN.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D10074; BAA00968.1; -.
                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            115
191
383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        116
                                                              POLG_HCVJ2
P27959;
                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                      POLG_HCVJ2
RESULT 1
                                                                        HIDD THE FEET THE PRINCE OF OCCUPANT TO THE PROPERTY OF THE PR
```

us-09-491-146a-26.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Envelope glycoprotein [Contains: Capsid protein C (Core protein E2 (GP86) (GP70) (NI)] (Fragment).
Hepatitis C virus (isolate HCV-476) (HCV).
Hepatitis C virus (isolate HCV-476) (HCV).
Hepatitis C virus (isolate HCV-476) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Abe K., Inchauspe G., Fullsawa K.;
"Genomic characterization and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-I- SUBBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                             (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                               Score 142; DB 1; Length 513;
Pred. No. 6.4e-13;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR002522; HCV_capsid.
Interpro; IPR002521; HCV_core.
Interpro; IPR002511; HCV_core.
Interpro; IPR002519; HCV_env.
Interpro; IPR002531; HCV_env.
Interpro; IPR002531; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_env; 1.
ProDom; PF01560; HCV_MS1; 1.
ProDom; PF01662; HCV_MS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                   943F31E3514CDEF3 CRC64;
                                                                                                                                 (GLCNAC. . .)
(GLCNAC. . .)
(GLCNAC. . .)
                           N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                       520 AA
                                                                                                                                                                                                                                                                                                                     1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                     PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-93019030; PubMed-1383400;
                                                                                                                                                                                                                   55704 MW;
                                                                                                                                                                                                                                                 91.0%;
llarity 92.9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D10688; BAA01530.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                448
513
513 AA;
                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-31643;
                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVH4_
ID POLG_HCVH4
AC Q01404;
                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                  NON_TER
SEQUENCE
                 TRANSMEM
                               CARBOHYD
                                                  CARBOHYD
                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 2
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lish-sib.ch).
                              MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1)] (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Abe K., Inchause G., Fullsawa K.;
"Genomic characterization and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic of non-A, non-B hepatitis in Japan.";
J. Gen. Virol. 73:2725-2729(1992).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN BAYELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus (isolate HCV-KF) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                            Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein; Coat protein; Envelope protein;
Transmembrane; Nonstructural protein.
PROTEIN (POTENTIAL)
PROTEIN (POTENTIAL).
ENVELOPE PROTEIN E (PC
                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                              AA135246CF20D525 CRC64;
                                                                                                                                                                                                                               Score 142; DB 1; Lenral No. 6.5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                            Pred. No. 6.5e
0; Mismatches
                                                                                                                                                                                                                                                                                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                        CAPSID |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-93019030; Pubmed-1383400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nterPro; IPR002522; HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
Pfam; PF01543; HCV_caps1d; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                               ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                        91.0%;
ilarity 92.9%;
Conservative C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; D10687; BAA01529.1; -. J01925; J01925.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                              56499
                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                              520 AA;
                                                                                                                                                                                                                                                                           Local Similarity
nes 26; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-31644;
                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVHK
Q01403;
                                                                                 CARBOHYD
                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                            Query Match
                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                     POLG_HCVHK
                                                                                                                                                                                                                                                                              Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
      211111111111
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                            ö
                           CAPSID PROPEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJORE ENVELOEB PROTEIN E (POTENTIAL).
NONSTRUCTUBAL PROTEIN NS1/E2 (POTENTIAL).
BY SIMILARITY.
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein); Matrix
E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
66nome polyprotein [Contains: Capaid protein C (Core protein); Matriprotein (Envelope protein M); Major envelope protein E; Nonstructura proteins (S1 and MS2] (Fragment).
Hepatitis C virus (isolate HC-15) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MNNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-9223023: PubMed-1314459;
Okamoto H., Kurai K., Okada S.I., Yamamoto K., Lizuka H.,
Okamoto H., Fuuda S., Tsuda F., Mishiro S.;
**Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                    Length 520;
   PROTEIN C
                                                                                                                                                                                                                                                                                                                                    Score 142; DB 1; Length 52
Pred. No. 6.5e-13;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                     MW; 1D2BD0A6FF27349B CRC64;
                  CELLULAR AMINOPEPTIDASE
   FROM CAPSID
                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     737 AA.
                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002531; HCV_NSI.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 D10075; BAA00969.1; -.
                                                                                                                                                                                                                                                                                                                                        91.08;
                                                                                                                                                                                                                                                                                                                                                      92.98;
                                                                                                                                                                                                                                                                                                     56476
                                                                                                                                                                                                                                                                                                                                                        Local Similarity 92.9
nes 26; Conservative
                                                                                                                                                                                                                                                                                                   520 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVJ5
ID POLG_HCVJ5
AC P27960;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genotypes."
                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                               CARBOHYD
                                                                                                       TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             8
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                   CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein); Matrix E; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genotypes.";
Virology 188.331.341(1992).

1- FUNCTION: THE SWALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC. SUGGESTHIG A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

1- SUBUNIT: THE VIRION OF THIS VIRIOS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-92230232; PubMed-1314459; Okamamoto K., Lizuka H., Okamato H., Kurai K., Okada S.I., Yamamoto K., Lizuka H., Canaka T., Fukuda S., Tsuda F., Mishiro S.; Fukula S., Fukula S., Fukula Comparatis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                             ΒY
                                                                                                                                             PROTEIN C
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Probom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Transmembrane; Nonstructural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core
protein (Envelope protein M); Major envelope protein
proteins NSI and NS2] (Fragment)
                                                                                                                                                                     CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%; Score 142; DB 1; L
92.9%; Pred. No. 9.5e-13;
Nomatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3AF699D82AD50131
                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus (isolate HC-77) (HCV). Viruses; ssRNA positive-strand viruses, no Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28
                                                                                                                                                                                                                                                                                                   POTENTIAL.
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81207 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJ7
P27961;
                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                        FRANSMEM
                                                                                                                                          INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJ7
AC DOLG_HCVJ7
AC DOLG_HCVJ7
AC DI AUG
DT 10-AUG
DT 10-AUG
DT 10-AUG
DE Protei
DE Protei
OS Viruse
OC Hepati
OX NCBLZI
NX NCBLZ

                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
ö
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-505 (Rel. 23, Created)
01-505 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 42, Last annotation update)
05-60-1992 (Rel. 42, Last annotation update)
05-60-1992 (Gorda) (GP32) (GP32
                                                                                                                                                                                                                                                                                                                             CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ./E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ. (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                    Nonstructural protein.

1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 142; DB 1; Length 737;
Pred. No. 9.5e-13;
); Mismatches 2; Indels
                                                                                                                                                                                                                                      Pfam; PF01560; HCV_NS1; 1.
ProDom; PD186062; HCV_NS1; 1.
Polyprotein; Glycoprotein; Coat protein; Envelope protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81691 MW; 67DFAE11854122F2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.0%; Scur.
92.9%; Pred. No. y...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                       InterPro; IPR00252; HCV_capsid.
InterPro; IPR00252; HCV_core.
InterPro; IPR00253; HCV_core.
InterPro; IPR00253; HCV_env.
InterPro; IPR00253; HCV_env.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
                                                                                                          EMBL; D10077; BAA00971.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity 92.9 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            737 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepacivirus.
NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLC_HCVBK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
                                                                                                                                                                                                                                                                                                 INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 6
POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
   q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X MEDLINE-98227846; PubMed-9568891;
X MEDLINE-98227846; PubMed-9568891;
Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M., Steinkuehler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.;
T. "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2-A resolution structure in a hexagonal crystal form.";
T. Protein Sci. 7:837-847(1998).
C. "FUNCTION: THE SWALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC. SUGGESTING A POSSIBLE MEBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLLE IN THE VIRAL RNA REPLICATION.
C. "CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral presence of precursor polyprotein, commonly with Asp or Giu in the P6 precursor polyprotein."
C. "CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
MEDLINE-91140698; PubMed-1847440;
Takamizawa A., Morl C., Fuke I., Manabe S., Murakami S., Fujita J.,
Onlshi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N., Moomaw E.W., Adachi T., Hostomska Z.; The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                     Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L. "Non-structural protein 3 of hepatitis C virus inhibits phosphorylation mediated by CANP-dependent protein kinase."; Eur. J. Blochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215 MEDLINE-97015088; Pubmed-8861916;
                                                                                                                                                                                                                                                                  SEQUENCE OF 1487-1500,
MEDLINE-96235224; PubMed-8647104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
                                                                                                                                                        from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PDB; 10UV; 26-JUN-00.
PDB; 80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002522;
InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A38465; GNWVTC.
1A1Q; 25-MAR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1JXP; 14-JAN-98
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-APR-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB; 1CSJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
```

```
STRAND
SEQUENCE
STRAND
STRAND
STRAND
STRAND
TURN
STRAND
HELIX
TURN
TURN
TURN
TURN
TURN
TURN
                                                                                                                                                                                STRAND
                                                                                                                                                                                           HELIX
 Q
                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FACE.

CELLULAR AMINOPEPTIDASE.

CAPSID PROTEIN C (POTENTIAL).

MATRIX PROTEIN (POTENTIAL).

MAJOR ENVELOPE PROTEIN E (POTENTIAL).

NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).

NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).

PROFESE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                       PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
CHARGE REL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
                                InterPro; IPR002166; HCV_RGRP.
InterPro; IPR001065; RRA_POI_DS_PS.
InterPro; IPR001065; RRA_POI_DS_PS.
InterPro; IPR001094; RRA_POI_PSVIr.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01559; HCV_env; 1.
Pfam; PF01560; HCV_NS2; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS5s; 1.
Pfam; PF0100999; Viral_RGRP; 1.
Pfam; PF00999; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE
ATP (PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645
2041
22077
22240
22240
22788
11035
11047
11063
                                                                                                                                                                                                                                                                                                           1081
1085
1087
1092
1094
1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1076
                                                                                                                                                                                                                                                                       3D-structure,
                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                  INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                             FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
HELIX
STRAND
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
TURN
                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
DORAN WAR AND CONTRACT OF THE CONTRACT OF THE
```

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21-); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV).
Virusee; SSNNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDILINE-91088550; PubMed-2175903;
Kato N., Hijlaka M., Oofsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Apanese patients with non-B hepatitis";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990);
                                                                                                                                                                                                                                                                                                                                                Score 142, DB 1; Length 3010;
Pred. No. 4.5e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                  327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                   91.0%;
                                                                                                                                                                                                                                                                                                                                                  Query Match 91.0 Best Local Similarity 92.9 Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
1112
1120
1120
1136
1136
11161
1161
1171
1174
11186
11186
11187
                                                                                                                                                                                                                                                                                                              3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11116;
11108
11120
11120
11120
11120
11162
11168
11175
11189
111980
111980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJA
```

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL)

(POTENTIAL) .

(POTENTIAL)

ö

ö

Length 3010;

us-09-491-146a-26.rsp

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last aenocation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP88) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99..); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE=98154321; PubMed=9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonuclectide: the crystal structure provides insights into the mode
of unwinding";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -I- FUNCTION: PROTESSE NS IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-IF FUNCTION: PROTESSE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-I- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Prince A.M.;
"Genomic structure of the human prototype strain H of hepatitis
"fenomic structure of the human prototype strain H of hepatitis
virus: comparison with American and Japanese isolates.";
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                   ; DB 1; Leus-
. 4.5e-12; 2; Indels
               N-LINKED (GLCNAC. . .) (POTENY 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-97331322; PubMed-9187654; YAO N. Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., "Structure of the hepatitis C virus RNA helicase domain."; Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                     91.0%; Score 142;
                                                                                                                                                                                                                                                                                                                                                                                                                        1 PKPORKPNRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                         Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         .
0
                                                                                                                                                                                                                                                 327017
                                                                                                                                                                                                                                                                                                                                   llarity 92,9%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Structure 6:89-100(1998)
                                                                     623
645
2041
2077
2240
2788
                                                                                                                                                                                                                                                    A.
                                                                                                                                                                                                                                                                                                                                   Local Similarity
nes 26; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepacivirus.
NCBI_TaxID=11108;
                                                                     623
645
2041
2077
2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (RNA) (N)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVH
P27958;
                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FT FT FT SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Glycoprotein; Envelope protein; Helicase, ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease
Transmembrane; Nonstructural protein; Hydrolase; Serine protease
Transmembrane; Nonstructural protein; Hydrolase; Serine protease
Transmembrane; Nonstructural REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.

CHAIN 115 CAPSID PROTEIN C POTENTIAL).
CHAIN 192 383 MAJOR ENVELOPE PROTEIN NSI (POTENTIAL).
TCHAIN 194 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
TCHAIN 1007 1615 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
TCHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
TCHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
   PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002522; HCV_capsid.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002531; HCV_capsid.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR001490; HCV_RAP.
InterPro; IPR001656; HCV_RAP.
InterPro; IPR001656; HCV_RAP.
InterPro; IPR001056; HA_POl_DS_PS.
InterPro; IPR001059; RNA_POl_DS_PS.
InterPro; IPR007099; RNA_POl_DS_PS.
InterPro; IPR007099; RNA_POl_DS_PS.
InterPro; IPR007099; HCV_CAPS.
InterPro; IPR007099; HCV_CAPS.
InterPro; IPR007099; HCV_CAPS.
InterPro; IPR007099; HCV_CAPS.
InterPro; IPR01538; HCV_NS1: InterPro; IPR01809; HCV_NS2: InterPro; IPR01006; HCV_NS4p; InterPro; IPR01006; HCV_NS4p; InterPro; IPR01001; HCV_NS4p; I
                                                                                                                                                                                                                                                                                                                 EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00271; hellcase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410; D
InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002519; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2013
3010
369
11083
11107
11165
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
12137
1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                          HSSP; P26663; 1JXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1083
1107
11165
1230
1316
1316
209
234
250
417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
```

Weber P.C.;

U

8888888888888888888

```
(POTENTIAL).
        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       327142 MW; 772CBB29CCD94753 CRC64;
        (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1635
1640
3011
                                                                                                                                                                                                                                           1224
1232
1233
1233
1247
1247
1272
1272
1273
1302
1312
1313
1323
1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1382
1389
1397
1410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1419
1432
1450
1456
1471
1480
1481
1489
1507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1343
1352
1360
1362
1368
1373
      CARBOHYD
CARBOHYD
CARBOHYD
                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
STRAND
STRAND
TURN
HELIX
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
STRAND
TURN
STRAND
TURN
STRAND
STRAND
TURN
STRAND
STRAND
TURN
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
TURN
STRAND
TURN
STRAND
TURN
STRAND
STRAND
TURN
STRAND
TURN
STRAND
TURN
HELIX
TURN
STRAND
TURN
TURN
HELIX
TURN
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
STRAND
STRAND
HELIX
SEQUENCE
                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transferase; RNA-directed RNA polymerase;
Envelope protein; Helicase; ATP-binding;
al protein; Hydrolase; Serine protease;
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
-!- PTW: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
-!- SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
-!- SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001512; HCV_capsid.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR00251; HCV_capsid.
InterPro; IPR00251; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR00166; HCV_NS5.
InterPro; IPR00266; HCV_RGP.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; INA_pol_DS_PS.
InterPro; IPR007094; INA_pol_DS_PS.
InterPro; IPR007094; INA_pol_PSvir.
Pfam; PP01540; HCV_capsid; 1.
Pfam; PP01539; HCV_capsid; 1.
Pfam; PP0150; HCV_NS1; 1.
Pfam; PP01006; HCV_NS1; 1.
Pfam; PP01006; HCV_NS3; 1.
Pfam; PP01001; HCV_NS4b; 1.
Pfam; PP01001; HCV_NS4b; 1.
Pfam; PP00271; helicase_C: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transmembrane; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; T
Core protein; Coat protein;
                                                                                                                                                                                                                                                   EMBL; M67463; AAA45534.1;
PIR; A36814; GNWVCH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
383
746
809
1026
11711
11711
11872
3011
369
1083
11083
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                    PDB; 1HEI; 25-NOV-98.
PDB; 1AIV; 16-FEB-99.
PDB; 1AIR; 17-JUN-98.
MEROPS; 529.001; -.
                                                                                                                                                                                                                                                                                                                                                                 TRANSFAC; T04155; -.
InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
RANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
POLG_HCVJ8
DORAN NEW WINDERS OF THE PRINCE OF THE PRINC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            144E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@iBb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WEDLINE-9204440; PubMed-1658196;
WASCIENCE TOWN NATURE
WAS GRANDED H., Okada S.-I., Suglyama Y., Kurai K., Lizuka H.,
Namoto H., Okada S.-I., Suglyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayumi M.,
Machedida Sequence of the genomic RNA of hepatitis C virus isolated
from a human carrier: comparison with reported isolates for conserved
if and divergent regions. 7.
Gen. Virol. 72:2697-2704(1991).
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAX A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAX A ROLE IN THE VIRAL RNA REPLICATION.
CHYDRITIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thi in pl and Ser or Ala in pl'.
C -- CATALITY ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotations (GPG)
29-FEB-2003 (Rel. 41, Last annotations (GPG)
20-FEB-2003 (NSI); Protein P7: Nonstructural protein NSA (P4P); Nonstructural protein NSA (P4P); Nonstructural protein NSA (P5P); Nonstructural protein NSSA (P6P); Nonstructural protein NSSB (P6E) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
20-FEB-2003 (Rel. 41, Last annotations (RCL) (RNA-directed RNA polymerase) (EC 2.7.7.48);
20-FEB-2003 (Rel. 41, Last annotations (RNA-directed RNA polymerase) (EC 2.7.7.48);
20-FEB-2003 (Rel. 41, Last annotations (RNA-directed RNA polymerase) (EC 2.7.7.48);
20-FEB-2003 (Rel. 41, Last annotations (RNA-directed RNA polymerase) (EC 2.7.7.48);
20-FEB-2003 (Rel. 41, Last annotations (RNA-directed RNA polymerase) (EC 2.7.7.48);
20-FEB-2003 (RRIA (RNA-directed RNA polymerase) (RRIA (RNA-directed RNA polym
                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                       ö
             91.0%; Score 142; DB 1; Length 3011;
llarity 92.9%; Pred. No. 4.5e-12;
Conservative 0; Mismatches 2; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                            1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                 HCV_caps1d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303. HSSP; P27958; JHEI. MREROPS; S.29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   interPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001490;
                Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                              POLG_HCVJ6
ID POLG_HCVJ6
AC P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                      02555586
025556
```

```
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2; · Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                 CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142; DB 1;
Pred. No. 4.6e-12;
0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23, Created)
23, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                            115
191
383
733
1010
1619
1866
2017
3033
369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 01-AUG-1992 (Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
```

a

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
           Genome polyprotein (Contains: Capsid protein ( (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP69) (GN31); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (Laolate HC-J8) (HCV).
                                                                                                                                                                                                                                                                                                                                                          JIONGOY 188:331-341(1992).

-I- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 posttion, Cys or Thr in Pl and Ser or Ala in Pl'.

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROPERIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                             MEDLINE-92230232; PubMed-1314459; Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka Vakuda S., Tsuda F., Mishiro S.; "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001490; HCV_NS4b.
InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_RGRP.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NSI.
HCV_NS3.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01560; HCV_NS1; 1. Pfam; PF01560; HCV_NS2; 1. Pfam; PF0207; HCV_NS3; 1. Pfam; PF01006; HCV_NS4s; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pr0Dom; PD186062; HCV_NS1; 1. SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                nterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01542; HCV_core; PF01539; HCV_env; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A40250; GNWVJ8.
; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THEI.
                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                NCBI_TaxID=11115;
                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
```

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 22, Last annotation update)
01-AUG-1992 (Rel. 22, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GPV0) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate I) (HCV)
                                             CELLULAR AMINOPEPTIDASE.
CARSID PROTEIN ( (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Halicase; APP-binding; Transmembrane; Nonstructural protein; Hadrolase; Serine protease.

INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                       RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
POTENTIAL.
                                                                                                                                                                              CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-91172826; PubMed-1848704;
Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 142; DB 1; Length 3033; Pred. No. 4.6e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                          N-LINKED (GLCNAC
                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                             (GICNAC
                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PKPORKTKRNTNRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      330177 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 92.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                             733
1010
1619
1866
2017
3033
3033
3033
1087
1111
1169
1241
1323
196
209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11104;
                                                                                                                                                                  347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCV1
                                                                                                                                                                                          ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                  TRANSMEM
                                                                                             CHAIN
CHAIN
CHAIN
                                                           CHAIN
                                                                                                                                              CHAIN
                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
POLG_HCV1
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
```

SIMILARITY). SIMILARITY). SIMILARITY).

(BY (BY (BY

(POTENTIAL)

(GLCNAC

POTENTIAL POTENTIAL POTENTIAL

(POTENTIAL) (POTENTIAL)

(GLCNAC

```
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                        Score 138; DB 1; Length sv. Pred. No. 1.7e-11;
                                                                                                                                                                                                                                                                                                                                                             65F8C9447FCE5AF9 CRC64;
                     POTENTIAL.
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                            2240
2240
2364
2789
327197 M
                                                                                                                                                                                                                                                                                                                                                                                     88.58;
                                                                                                                                                                                                                                                                                                                                                                                                 88.38;
                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 89.3 tes 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                           2041 204
2077 207
2240 224
2364 236
2789 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCVTW
                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                           LSITE
BIND
                                                                                                 CARBOHYD
                                                                                                                                                                 CARBOHYD
                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MAJNEX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
Bradley D.W., Kuo G., Houghton M.;
Genetic organization and diversity of the hepatitis C virus.";
Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).
- Frunction: THE SMALL PROFEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELEATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postion, Cys or Thr in P1 and Ser or Ala in P1'.
- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                          (RNA)(N).

SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein; Envelope protein; Helicase; ATP-binding;
structural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                           SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001400; HCV_NS4.
InterPro; IPR001400; HCV_NS4.
InterPro; IPR001560; HCV_NS5a.
InterPro; IPR001560; Helicase_C.
InterPro; IPR001050; Helicase_C.
InterPro; IPR007095; RNA_POl_DS_PS;
InterPro; IPR007095; RNA_POl_DS_PS;
InterPro; IPR007094; RNA_POl_PSvIr.
Pfam; PF01543; HCV_COTE; I.
                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522; HCV_caps1d.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nterPro; IPR002518; HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transmembrane; Nonstructural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     interPro; IPR002531; HCV_NS1
                                                                                                                                                                                                                                                                                                                                      EMBL; M62321; AAA45676.1; --
PIR; A39166; GNWVC3.
PDB; IA1V; 16-FBB-99.
PDB; 1HEI; 25-NOV-98.
MEROPS; S29.001; --
MEROPS; U39.001; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00271; helicase_C; 1
Pfam; PF00998; Viral_RdRP; 1
ProDom; PD186062; HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD InterPro; IPR002522; HCV_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01006; HCV_NS4a;
PF01001; HCV_NS4b;
PF01506; HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01538; HCV_NS2; ]
PF02907; HCV_NS3; ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01560; HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
```

ö

Gaps

; 0

Length 3011;

(POTENTIAL).
(POTENTIAL).
(POTENTIAL).

(GLCNAC.

(GLCNAC

POTENTIAL (POTENTIAL) POTENTIAL

POTENTIAL) POTENTIAL POTENTIAL

(GLCNAC.

(GLCNAC

```
01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
60-APR-1993 (Rel. 42, Last annotation update)
60-600me polyprotein (Contains: Capsid protein C (Core protein) (P22);
60-603 (GP70) (NG1); Protein P7; Nonstructural protein NS2 (P21)
60-603 (GP70) (NG1); Protease/helicase NS3 (P70) (Hepacivirin)
60-603 (GP70); Nonstructural protein NS5A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P65); No
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92230206; PubMed=1314449;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
The Talwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA.";
Virology 188:102-113(1992).
I-FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
I-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
I-CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (knA)(w).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PRT; 3010 AA
```

Gaps

ö

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE=92295714; PubMed=1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata M., Ishimura Y., Shimotohno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VITUE RES. 23:39-53(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
    POTENTIAL)
                                                                                                                (POTENTIAL)
                                                                                                                              (POTENTIAL)
                                                                                                                                          (POTENTIAL)
                    POTENTIAL)
                                POTENTIAL
                                                                                                                                                          (POTENTIAL
                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                   MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                             Score 135; DB 1; L
Pred, No. 4.6e-11;
0; Mismatches 2;
                                           (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                    (GLCNAC.
                                                                                                              (GLCNAC.
                                                                                                                                           (GLCNAC
                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                      3010 AA
                                                                                                                                                                                                                                          0; Mismatches
                                                                                                                                                                                                                                                                      2 KPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                    6 KPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                N-LINKED
86.5%;
92.6%;
                                                                                                                                                                                                                        Local Similarity 92.6 les 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTEIN C AND MRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=31642;
                                                                                   623
645
2041
2077
2240
2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJT
Q00269;
                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                Query Match
                                             CARBOHYD
                                                           CARBOHYD
                                                                        CARBOHYD
                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                       POLG_HCVJT
                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                          RESULT 13
    g
                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTESAE/HELICASE NS3 (POTENTIAL).
PROTESAE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fransmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BY SIMILARITY).
(BY SIMILARITY).
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE RELAY SYSTEM (1)
CHARGE RELAY SYSTEM (1)
CHARGE RELAY SYSTEM (1)
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                          | IPR002531; HCV_NS1.
| IPR002531; HCV_NS1.
| IPR004109; HCV_NS3.
| IPR00145; HCV_NS4a.
| IPR001490; HCV_NS4b.
| IPR002868; HCV_NS5a.
| IPR002166; HCV_RGAP.
| IPR0071095; RNA_DO1_DS_PS.
                                                                                                                                                                    EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                    HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF01538; HCV_NS2; 1.
Pfam; PF01507; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NSAb; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
Probom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01543; HCV_caps1d;
PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1006
1615
1862
2013
3010
3010
11003
11165
1237
1319
1319
1319
233
233
4417
423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; DEXDC;
                                                                                                                                                                                                                                                       InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                 InterPro; IPR002519;
                                                                                                                                                                                              1NS3; 08-APR-98
                                                                                                                                                                                                                                                                   IPR002522;
                                                                                                                                                                                   A40244; GNWVTW
                                                                                                                                                                                                                          MEROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     347
1083
1107
11165
11230
1316
196
233
233
234
250
417
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01539; PF01560; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                          MEROPS; U39.
                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                            nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      fam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
STRAIN=New Zealand; TISSUE-Brain, and Skeletal muscle;
MEDLINE-96161994; PubMed-8576224;
CSOTTOS C., Zolnierowicz S., Bako E., Durbin S.D., Depaoli-Roach A.A.;
"High complexity in the expression of the B' subunit of protein phosphatase 2A0. Evidence for the existence of at least seven novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBBUILT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME, COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa CONSTANT REGULATORY SUBUNIT (PR65 or SUBUNIT A), THAT ASSOCIATES WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS BY THE CAPE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNITS AND CELL SIGNALING MOLECULES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein; Repeat. 8 X 2 AA APPROXIMATE TANDEM REPEATS OF Q- \theta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1soforms.";
J. Biol. Chem. 271:2578-2588(1996).
-:- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
                                                                                                                                                                                                                                                                                                    DASD_RABIT STANDARD; PRT; 586 AA.

Q28653; Q28655;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PP2A, B subunit, BY delta isoform) (PP2A, B subunit, PK61 delta isoform) (PP2A, B subunit, R5 delta isoform) (PP2A, B subunit, R6 delta isoform)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAUTION: NOMENCLATURE USED IN REF.1 REFERS TO PP2A B SUBUNIT B' GAMMA ISOFORM, WHICH IS CITED AS PP2A B SUBUNIT DELTA-PR61 ISOFORM IN LATER PUBLICATIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -i- SUBCELLUIAR LOCATION: Nuclear (By similarity).
-i- TISSUE SPECIFICITY: Highly expressed in brain.
-i- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit
                                                                                              ö
                                            Score 134; DB 1; Length 3010;
Pred. No. 6.4e-11;
0; Mismatches 3; Indels
326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, But
Mammalia, Eutheria, Lagomorpha, Leporidae, Oryctolagús.
                                                                                                                                          1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                   Oryctolagus cuniculus (Rabbit).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; U38193; AAC48532.1; -.
                                            85.98;
                                                                  Best_Local Similarity 89.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U38195; AAC48534.1;
InterPro; IPR002554; B56.
Pfam; PF01603; B56; 1.
Multigene family; Nuclear p
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9986;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                507
SEQUENCE
                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                          RESULT 14.
2A5D_RABIT
                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDAGE.
CAPSID PROTEIN ( POTENTIAL).
MATRY PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
CHARGE RELAY SYSTEM (CHARGE RELAY SYSTEM (CHARGE RELAY SYSTEM (CHARGE RELAY SYSTEM (ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR007095; RNA_pol_DS_PS.
                                                                                                                                                            HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01543; HCV_copsid; 1. Pfam; PF01542; HCV_core; 1. Pfam; PF01539; HCV_core; 1. Pfam; PF01539; HCV_nov; 1. Pfam; PF01569; HCV_NS1; 1. Pfam; PF01001; HCV_NS3; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01001; HCV_NS4b; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00271; helicase_C; 1. Pfam; PF00271; PF0
                                                                                                                                                                                                                                                                           HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                      MEACEPRO; U.Y.C.Y.
InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_capsi
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SM00487; DEXDC;
         PIR; A45573; A45573.
PDB; IA1Q; 25-MAR-98.
PDB; IJXP; 14-JAN-98.
MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                  IPR002518;
IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                   IPR002868;
                                                                                                                                                                                                                                                                                                                          IPR001490;
                                                                                                            U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3D-structure
INIT_MET
                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProDom; F
SMART; SM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
```

SO

q

à

```
5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDINE-2238257; PubMed-12477932;
MEDINE-2238257; PubMed-12477932;
MEDINE-2238257; PubMed-12477932;
MEDINE-2238257; PubMed-12477932;
MALSCHIE R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T. L., Scheetz T.E., Bronstein M.J., Uddin T.B., Toshlyuki S., Carninci P., Prange C., Andrews M., McErnan K.J., Malek J.A., Gunarathe P.H., McEwan P.J., McKernan K.J., Malek J.A., Gunarathe P.H., Allalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Willalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Allalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Millalon M., Touchman J.W., Green E.D., Dickson M.C., Allakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Allakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Allakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             subunits
PP2A to
                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM DELTA-2), AND SEQUENCE OF 501-508; 550-559; 573-560 AND 584-601 (DELTA-1).
TISSUE-Bone marrow, and Brain cortex;
TISSUE-BONE marrow, and Brain cortex;
Tanabe O., Nagase T., Murakami T., Nozaki H., Usui H., Nishito Y., Hayashi H., Kagamiyama H., Takeda M.;
"Molecular cloning of a 74-kba regulatory subunit (B', or delta) of human protein phosphatase 2A.";
                                                                                                                                                                                                                                                                              014738; 000494; 000696; 015171; 01.00 MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit, delta isoform (PPAA, B subunit, B' delta isoform) (PP2A, B subunit, B' delta isoform) (PP2A, B subunit, R5 delta isoform) (PP2A, B subunit, R5 delta isoform) (PP2A, B subunit, BS delta isoform) (PP2A, B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human). '
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE-Brain cortex;
MEDLINE-97324098; PubMed-9180267;
Tanabe O., Gomez G.A., Mishito Y., Usui H., Takeda M.;
Tanabe I., Gomez G.A., Mishito Y., Usui H., Takeda M.;
"Molecular heterogeneity of the cDNA encoding a 74-kDa regulatory subunit (B', or delta) of human protein phosphatase 2a.";
                                                                              11;
                                        Length 586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE-Fetal brain;
MEDLINE-96355607; PubMed-8703017;
McCright B., Rivers A.M., Audlin S., Virshup D.M.;
The B56 family of protein phosphatase 2A (PP2A) regulatory encodes differentiation-induced phosphoproteins that target both nucleus and cytoplasm.";
                                                                              Indels
E149A309CDDA7495 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-3).
                                                                                                                                          1:+| :| :|:|| 30 PQPQSQP-PSSNKRPSNSTPPPTQLSKIKYSGGPQIV 65
                                                                                                                   1 PKPQRKPNRNTNRRPQD------VKFPGGGQIV 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. (ISOFORMS DELTA-1 AND DELTA-2)
                                                                            5,
                                      DB 1;
                                    Score 55.5; DB Pred. No. 2.2; 8; Mismatches
                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                 602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. (ISOFORM DELTA-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 h nucleus and cytoplasm.";
Biol. Chem. 271:22081-22089(1996).
                                                                                                                                                                                                                                                               PRT;
                                                                            .;
8
68090 MW;
                                    35.6%;
illarity 35.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEBS Lett. 379:107-111(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEBS Lett. 408:52-56(1997).
586 AA;
                                      Query Match
Best Local Similarity
Matches 13; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                    2A5D_HUMAN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                     PPP2R5D
                                                                                                                                                                                                                                        2A5D_HUMAN
ID 2A5D_H
AC Q14738
DT 30-MAY
DT 15-SEP
DE Serine
```

RTH REPRESENTED BY THE REPRESENT BY THE

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MIM, 50106634; -.
GO; 0005634; C:nucleus; TAS.
GO; GO:0008601; F:protein phosphatase type'2A, regulator acti. . .; TAS.
GO; GO:0007399; P:neurogenesis; TAS.
InterPro; IPR002554; B56.
                                                                                                                                                                                                                                                          SUBBUILT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
COMPOSED OF A 36 KDA CATALYTIC SUBUNIT (SUBUNIT C). AND A 65 KDA
CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATES
WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
B (THE R2/B/PR55/B55, R3/B''/PR72/PR130/PR59 AND R5/B'/B56
FAMILIES), THE 48 KDA VARIABLE REGULATORY SUBUNITY
AND CELL SIGNALING MOLECULES.
SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR IN INTERPHASE,
NUCLEAR DUBLING MILESIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SH3 BINDING, CLASS I (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schorch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: THE B REGULATORY SUBUNT MIGHT MODULATE SUBSTRATE SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ISOId-014738-3; Sequence=VSP_005110;
--- TISSUE SPECIFICITY: ISOFORM DELTA-2 IS WIDELY EXPRESSED. ISOFORM DELTA-1 IS HIGHLY EXPRESSED IN BRAIN.
---- INDUCTION: BY RETINOIC ACID; IN NEUROBLASTOMA CELL LINES.
---- PTM: AT LEAST ISOFORM DELTA-1 IS PHOSPHORYLATED ON SERINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8 X 2 AA APPROXIMATE TANDEM REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -1- SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein; Phosphorylation; Alternative splicing; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Missing (in isoform Delta-3)./FTId=VSP_005110.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Missing (in isoform Delta-2).
/FIId-vSP_005111.
F15F71AF4E565387 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Event-Alternative splicing; Named 1soforms-3;
Name-Delta-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 53.5; DI
Pred. No. 4.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IsoId=014738-2; Sequence=VSP_005111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IsoId=014738-1; Sequence-Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Missing (in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AB000634; BAA20381.1; -. AB000635; BAA20382.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D78360; BAA11372.1; -. BC001095; AAH01095.1; -. BC001175; AAH01175.1; -. BC010692; AAH10692.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; L76702; AAB69751.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              34.3%;
31.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ....ew; HGNC:9312; PPP2R5D.
GK; Q14738; ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALTERNATIVE PRODUCTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         530
565
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01603; B56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Multigene family.
DOMAIN 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Name - Delta - 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Name-Delta-3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                        COMPARTMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESIDUES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
```

Matches 13; Conservative 8; Mismatches 6; Indels 15; Gaps

3

1 PKPQRKPNR-----NTNRRPQD------VKFPGGGQIV 27 |:|| :| | || || 40 PQPQPQAQSQPPSSNKRPSNSTPPPTQLSKIKYSGGPQIV 81

ò qq Search completed: August 7, 2003, 11:19:59 Job time: 5.90909 secs

```
94.9
94.9
94.9
9.0
 17
18
19
20
                                                                                                   August 7, 2003, 11:05:41; Search time 25.6364 Seconds (without alignments) 281.845 Million cell updates/sec
                                                                                                                                                                                                                                                                                                    830525
              GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                   tal number of hits satisfying chosen parameters:
                                                                                                                                                                US-09-491-146A-26
156
1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                     830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                          OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_plant:*
sp_rodent:*
sp_virus:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_archea:*
sp_bacteria:*
sp_fung1:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                         SPTREMBL_23:*
                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                        Searched:
                                                                                                       Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

4		æ (					
No.	Score	Query	Duery Match Length DB	DB	ID	Description	uo
-	150	96.2	191	12	Q8V7S2	08v7s2 he	hepatitis c
7	150	96.2	191	12	Q8V7T0		hepatitis c
e	150	96.2	191	12	Q8V7S4		epatitis c
4	148	94.9	119	12	081291		hepatitis c
S	148	94.9	191	12	Q8JP76	08jp76 he	epatitis c
9	148	94.9	191	12	Q8V7XB	Q8v7x8 he	hepatitis c
7	148	94.9	191	12	Q8V7V4	Q8v7v4 he	epatitis c
8	148	94.9	191	12	Q8V7X5	Q8v7x5 he	epatitis c
σ	148	94.9	191	12	Q8V7Q3	Q8v7q3 he	epatitis c
10	148	94.9	191	12	Q8V7K4	08v7k4 he	hepatitis c
11	148	94.9	191	12	Q8V7M4	Q8v7m4 he	epatitis c
12	148	94.9	191	12	Q8V7L1	Q8v711 he	hepatitis c
13	148	94.9	191	12	Q8V7Y3	Q8v7y3 he	hepatitis c
14	148	94.9	191	12	Q8V7N6		hepatitis c
15	148	94.9	191	12	Q8V7Y0		hepatitis c
16	148	6 7 6	191	13	ORV7K9		C Stities

17   148   94.9   191   12 087746   0897   188   148   94.9   191   12 087746   0897   188   148   94.9   191   12 087777   089778   089778   188   94.9   191   12 087777   089778   089778   089778   148   94.9   191   12 087777   089778   0897	hepatitis hepatitis hepatitis hepatitis hepatitis	hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis	08778 hepatitis c 08778 hepatitis c 088772 hepatitis c 06831 hepatitis c 06831 hepatitis c 06830 hepatitis c 06830 hepatitis c	r8 hepatitis r7 hepatitis r7 hepatitis 12 hepatitis 82 hepatitis 08 hepatitis	MENTS	191 AA.	uence update) otation update) ragment).	es, no DNA stage; Flaviviridae;	ancerous and non-cancerous hepatic	EREU EINS IPLEN	CAB3695EF34 CRC64;	50; DB 12; Length 191; o. 7e-14: atches 1; Indels 0; Gaps 0;	G 28
17 148 94.9 191 12 18 148 94.9 191 12 20 148 94.9 191 12 21 148 94.9 191 12 22 148 94.9 191 12 23 148 94.9 191 12 24 148 94.9 191 12 25 148 94.9 191 12 26 148 94.9 191 12 27 148 94.9 191 12 28 148 94.9 191 12 28 148 94.9 191 12 28 148 94.9 191 12 29 145 92.9 191 12 31 145 92.9 191 12 32 144 92.3 191 12 33 144 92.3 191 12 34 142 91.0 66 12 35 144 92.3 191 12 36 142 91.0 61 12 37 142 91.0 61 12 38 142 91.0 61 12 39 142 91.0 61 12 30 144 92.3 190 12 31 145 92.9 191 12 31 145 92.9 191 12 32 144 92.3 190 12 34 142 91.0 61 12 35 144 92.3 190 12 36 145 92.9 191 12 37 142 91.0 61 12 38 144 92.3 190 12 39 145 92.9 191 12 30 145 92.9 191 12 30 145 92.9 191 12 30 145 92.9 191 12 31 145 92.9 191 12 31 145 92.9 191 12 31 145 92.9 191 12 32 144 92.3 190 12 33 144 92.3 190 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 61 12 34 142 91.0 74 12 35 144 92.3 191 12 36 144 92.3 191 12 37 142 91.0 74 12 38 144 92.3 191 12 39 145 144 92.3 191 12 30 145 144 92.3 191 12 30 145 144 92.3 191 12 31 145 144 92.3 191 12 31 145 144 92.3 191 12 31 145 147 147 147 147 147 147 147 147 147 147	Q8V7W1 Q8V7W6 Q8V7N7 Q8V7X7 Q8V7X9	08V7M0 08V7K8 08JW46 08V7L7 08V7M7 08V7W3	Q8V7W8 Q8V7W5 Q8V7Y2 Q68520 Q68311 Q68306 Q68309	Q8JYR8 Q8JYR5 Q8JYR7 Q8JYR6 Q68712 Q68682 Q68682	ALIGN	PRT;	Created) Last seq Last ann	ind virus	ies in c	EMBL/Gen HIS VIRU IE ENVELO IIN E. TH SIMILARIT	7; 64151	Score 15 Pred. No 0; Misma	FPGGGQIV
17 148 94.9 191 18 148 94.9 191 20 148 94.9 191 21 148 94.9 191 22 148 94.9 191 23 148 94.9 191 24 148 94.9 191 25 148 94.9 191 26 148 94.9 191 27 148 94.9 191 28 145 92.9 191 29 145 92.9 191 30 145 92.9 191 31 145 92.9 191 31 145 92.9 191 32 144 92.3 191 33 142 91.0 60 40 142 91.0 61 41 142 91.0 61 41 142 91.0 61 42 142 91.0 61 44 142 91.0 61 44 142 91.0 61 45 92.9 191 46 142 91.0 61 47 142 91.0 61 48 142 91.0 61 49 142 91.0 61 40 142 91.0 61 41 142 91.0 61 42 142 91.0 61 44 142 91.0 61 44 142 91.0 61 45 142 91.0 61 46 142 91.0 61 47 142 91.0 61 48 142 91.0 61 49 191 40 142 91.0 61 41 142 91.0 61 41 142 91.0 61 42 142 91.0 61 44 142 91.0 61 4	21 22 21 22 21 22 22 22 22 22 22 22 22 2	222222	22222222	2222222			20, 20, 22, 1yprc	-stra	tspec	the OF TEROTE SOLL; V.cor	40 M	2. 4. 0. 8 %;	PQDVK
TESULT 1  TESULT	თთთთთთ	თთთთთთ	000004446	61 62 7 7 7 7 7 7 7 7 7 7		MINARY	BLrel. BLrel. BLrel.	sitive	is quas.	VIRION VIRION SINVELOPI GLYCOI MRNA MRNA MRNA MRNA MRNA MRNA MRNA MRNA	191	96 y 96 rvative	RNTNRR!
TESULT 1  RESULT 1  ORVYS2  TO 1042 99  24 148 99  25 148 99  26 148 99  27 148 99  28 148 99  29 148 99  20 148 99  21 148 99  22 148 99  23 148 99  24 148 99  25 148 99  26 148 99  27 142 99  28 145 99  29 145 99  31 145 99  32 144 99  33 144 99  34 142 99  35 142 99  36 142 99  37 142 99  38 142 99  39 142 99  30 142 99  31 142 99  32 144 99  33 144 99  34 142 99  36 142 99  37 142 99  38 142 99  38 142 99  39 142 99  30 142 99  31 142 99  32 144 99  33 144 99  34 142 99  36 142 99  37 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  38 142 99  39 142 99  38 142 99  38 142 99  38 142 99  38 142 99  39 142 99  38 1142 99  38 1142 99  39 142 99  38 1142 99  38 1142 99  39 142 99  39 142 99  30 144 99  30 142 99  30 144 90  30 144 90  30 144 90  30 144 90  30 144 90  30 144 90  30 144 90  30 144 90  30 144 90				0000000		ELI	Se de la constant de	rus N pc (03;	N.A.	AND	₹ 2	rit	Ж. —
RESULT 1 0897782 01-0442 023 1148 024 1148 025 1148 026 1148 027 1148 0287782 03 1142 03 1142 03 1142 03 1142 03 1142 03 1142 03 1142 03 1142 03 1142 04 0 1142 04 0 1142 04 0 1142 05 0 142 06 0 142 07 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0	999999	9999999	2000000000	66888888		ď	5553	C vj 8RNA 8.	ROM er;	CMAN N C DTE DTE DTE DPR DPR DPR DPR DPR DPR DPR DPR DPR DPR	191	nila C	2007 
NO N	444444	444444	* 4 4 4 4 4 4 4	11111111111111111111111111111111111111		1 V7S2 V7S3.	-MAR-200; -MAR-200; -OCT-200; re prote	kE. patitis ( ruses; se pacivirus	) QUENCE FI SSUE=Live to N.; epatitis	bmitted - SUBUNI: LIPOPR( PROTEII PROTII	Lyprotell N_TER QUENCE	Match Local Sines es 27;	ш —
	11 118 20 21	222222 24222 2422 2422 2422 2422		W W 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4		RESULT 28V7S2 ID Q8	00000	OS A CO	RP SER	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	FT PO	Query Best Match	λά

음

```
Pfam; PF01542; HCV_core; 1.
                                                                            191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                       Polyprotein.
NON_TER 191
SEQUENCE 191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q8JP76;
                                                                                                                                                                                                                                                                                                                                                                                                        081291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08JP76
                                                                                                                                                                                                                                                                                                                                                      RESULT 4
Q81291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08JP76
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AB062194; BAB83388 1; -
InterPro; IPR002522; HCV_capsid.

InterPro; IPR002521; HCV_capsid.

Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-1- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 150; DB 12; Length 191;
Pred. No. 7e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191 191
191 AA; 20740 MW; 64151CA83695EF34 CRC64;
                                                                                                                                                                                     ou man-zouz (TrEMBLrel. 20, Last sequence update) 01-027-2002 (TrEMBLrel. 22, Last annotation update) core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Mismatches
PKPQRKPKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL, AB062200; BAB83394 1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             th 96.2%;
I Similarity 96.4%;
27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAR-2002 (TrEMBLrel. 20,
01-MAR-2002 (TrEMBLrel. 20,
01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                      01-MAR-2002 (TrEMBLrel. 20,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Liver;
Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8V7S4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8V7S4
                                                                       RESULT 2 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (089770 (08970 (08970 (08970 (08970 (08970 (08970 (08970 (08970 (08970 (08970 (08970 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08V7S4
ID 0
```

g

ò

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS5D region sequences.";
J. Clin. Microbiol. 32:3049-3051(1994).
I. Clin. Microbiol. 32:3049-3051(1994).
I. SUBUNIT THE VIRION OF THIS VIRIOS IS A NUCLEOCAPSID COVERD BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; D26383; BAA05399.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyak S.J., Miller K., McArdle S., Gale M.J. Jr., Tencever B., Hiscott J., Gretch D.R.; "Quasispecies Dependent Activation of the Innate Interferon System By
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hotta H., Handajani R., Ingelusida M., Soemarto W., Doi H., Miyajima H., Homma M.; "Subtype analysis of hepatitis C virus in Indonesia on the basis of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                       Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94.9%; Score 148; DB 12; Length 119; 96.4%; Pred. No. 8.2e-14; 1.ve 0; Mismatches 1; Indels
640D7CAB2765EF34 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             119 AA; 13510 MW; F6836FBFA5E5EB2A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                    Score 150; DB 12;
Pred. No. 7e-14;
0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                            119 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191 AA.
                                                                                                                                                            1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                         5 PKPQRKPKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
Core protein (Genome polyprotein).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Td-34/92;
MEDLINE-95189942; PubMed-7883898;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro, IPR002522; HCV_capsid.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein (Fragment). Hepatitis C virus.
20744 MW;
                                                    Query Match
Best Local Similarity 96.2%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 96.4 ies 27; Conservative
                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
```

ô

```
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flavlviridae;
                                                                                                                                                                                                     Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROFEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA GLY SIMILARITY).

EMBL, AB062025; BAB83364.1; -.

InterPro; IPR002521; HCV_capsid.

Pfam; PF01543; HCV_capsid, 1.

Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (MAY-2001) to the EMBL/GenBank/DD3J databases.

-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROPEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROFIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062004; BAB83343.1; -
InterPro; IPR002521; HCV_capsid.

InterPro; IPR002521; HCV_capsid.

Pfam; PF01543; HCV_capsid.

Pfam; PF01542; HCV_core; 1.
    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                 Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191 191 20728 MW; 23C8ABC1783B849A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER 191 191
SEQUENCE 191 AA; 20694 MW; 2F4BD80E7BFE8E9A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2002 (TrEMBLrel. 20, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11arity 96.4%; Score 148; DB 12; Observative 0; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 148; DB 12;
Pred. No. 1.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94.98;
96.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2002 (TrEMBLrel. 20,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 96.4
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
les 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rissue-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                    TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       lesions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                         lesions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08V7X5;
                                                                                                                                            Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q8V7X5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
       STATE TO COURT TANK SERVICE SOLUTION OF THE SERVICE SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The Hepatitis C Virus Core Protein.";
Submitted (JuL-2002) to the EMBL/GenBank/DDBJ databases.
1- SUBUNIT: THE VIRION OF THIS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
EMBL, AR531649, AAM95639.1; -.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                                                                                                                         Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13; 1ve 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                                                                                                                                              191 AA; 20797 MW; 3632F808598A348A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment)
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                       Score 148; DB 12;
Pred. No. 1.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191 AA
                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-2002 (TrEMBLrel, 20, Created)
                                                                                                                                                                                                                                                                                                                            94.98;
                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 96.4%;
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 94.99
Best Local Similarity 96.43
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IISSUE-Liver;
                                                                                                                                                                                                                                                              Polyprotein.
SEQUENCE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q8V7V4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   987789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q8V7V4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08V7V4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
         ò
```

Gaps

ö

ö

Gaps

```
SEQUENCE
                                                                                                                                   08V7M4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08V7L1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                 ò
                                                                                    g
                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                        "Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                         Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNT: THE .VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSIGTS OF TWO PROTEINS: PROTEIN M BUYELOPE THE ENVELOPE CONSIGTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
InterPro: IPRO02521; HCV_capsid.
InterPro: IPRO02521; HCV_capsid.
InterPro: IPRO02521; HCV_core.
Pfam: PF01543; HCV_capsid: 1.
                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                         Score 148; DB 12; Length 191;
Pred. No. 1.4e-13;
0; Mismatches 1; Indels 0
                                                                                                                                                                                                                                                                                                Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20794 MW; 36DD3CBAE63B8532 CRC64;
                                      01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
CORE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-2002 (TrEMBLrel. 20, Created)
1-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                      191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA.
                                                                                                                                                                                                                                                                                                                                                                                           S PKPQRKTNRNTNRRPQDVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                         94.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAR-2002 (TrEMBLrel, 20,
                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 96.45
Matches 27, Conservative
                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                   TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein.
NON_TER
                                                                                                                                                                                    lesions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kato N.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        lesions.
                                                                                                                                                               Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     08V7K4
RESULT 9
Q8V7Q3
                      g
                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                      144A1
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                        VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROFEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                            Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13; tive 0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        lesions.";
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
191 AA; 20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191
20746 MW; 23DB0B26D33B849A CRC64;
                                                                                                                                                                                                                                                                                   01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
                                  94.9%; Score 148; DB 12;
96.4%; Pred. No. 1.4e-13;
11ve 0; Mismatches 1;
                                                                                                                                                                                                                                                    191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA.
                                                                                                             1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                   5 PKPQRKTNRNTNRRPQDVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILLI FILLIFILIFILIFILIFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AB062250; BAB83444.1; -...
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_capsid; 1...
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 96.4%;
nes 27; Conservative
                                Query Match
Best Local Similarity 96.44
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               191 1
191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              lesions.";
Submitted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kato N.;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (MAY-2001) to the EWBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB062263; BAB83457.1; -
InterPro; IPR002521; HCV_capsid.
LinterPro; IPR002521; HCV_capsid.
FinerPro; IPR002521; HCV_capsid.
FinerPro; IPR01542; HCV_capsid.
FinerPro; IPR01542; HCV_capsid.
                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Hepatitis C virus quasispecies in cancerous and non-cancerous
                                                                                                                                                                                                                    94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13; tive 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94.9%; Score 148; DB 12; Length 191; 96.4%; Pred. No. 1.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                    Polyprotein.
NON_TER 191 191
SEQUENCE 191 AA; 20730 MW; 23C8A9ACE82A749A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2002 (TTEMBLrel. 20, Created)
01-MAR-2002 (TTEMBLrel. 20, Last sequence update)
01-OCT-2002 (TTEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
Core protein (Genome polyprotein) (Fragment).
CORE.
                                                                                                                                                                                                                                                                                                                                                                                                     191 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                        PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                     1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROTEIN M AND GIYCOPROTEIN E. THE N
PROTEIN C AND MRNA (BY SIMILARITY).
EMBL; AB061996; BABB3335.1; -.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2002 (TrEMBLrel. 20, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 96.41
Matches 27; Conservative
                                                                                                                                                                                                                                     Best Local Similarity 96.4 Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Polyprotein.
NON_TER 19
SEQUENCE 19:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            lesions."
                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8V7N6;
Q8V7N6;
                                                                                                                                                                                                                                                                                                                                                                                                       Q8V7Y3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kato N.
                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
Q8V7Y3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                                                                                         ö
   S T W D D R R B B R R B B F S
                                                                                                                                                                                                                                                                                       ò
```

```
"Hepatitis C virus quasispecies in cancerous and non-cancerous hepatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BY A
                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (MAX-2001) to the EMBL/GenBank/DDBJ databases.

-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AB061999; BAB83338.1; -
INTEPPO; PRR002522; HCV_capsid.

InterPro; PRR002522; HCV_capsid.

InterPro; PRR002521; HCV_core.

Pfam; PF01542; HCV_core; 1.
                                                                                                                               lesions.";
Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus quasispecies in cancerous and non-cancerous lesions. ";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            y Match 94.9%; Score 148; DB 12; Length 191; Local Similarity 96.4%; Pred. No. 1.4e-13; hes 27; Conservative 0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                        Length 191;
                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                        191
20730 MW; 23C8A9ACE82A749A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               191 191
191 Aa; 20718 MW; 23C8A9ACE83B849A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAR-2002 (TrEMBLrel. 20, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Core protein (Genome polyprotein) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                      Score 148; DB 12;
Pred. No. 1.4e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191 AA
                                                                                                                                                                                                                                                                                                                                                                      Query Match
94.9%; Score 148; DE
Best Local Similarity 96.4%; Pred. No. 1.46
Matches 27; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PKPQRKTNRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last seq
01-0CT-2002 (TrEMBLrel. 22, Last ann
                                                                                                                                                                                                                         EMBL; AB062238; BAB83432.1; -...
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                        191 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                              SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
                                                                               TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Liver;
                                                                                                                                                                                                                                                                                                         Polyprotein.
NON_TER 1
SEQUENCE 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                              lesions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kato N.;
                                                                                              Kato N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8V7Y0
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  D.
```

Search completed: August 7, 2003, 11:18:59 Job time : 26.6364 secs

```
Sequence 6, Appli
Sequence 36, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 23, Appli
Sequence 26, Appli
Sequence 3, Appli
Sequence 26, Appli
Sequence 19, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 118, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 26, Appl
                                                                                                                                                                          August 7, 2003, 11:07:41; Search time 10.5455 Seconds (without alignments) 112.343 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 8
Sequence 8
Sequence 8
Sequence 2
Sequence 4
Sequence 6
Sequence 6
Sequence 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1: /cgn2_6/ptodate/1/laa/5A_COMB.pep:*
2: /cgn2_6/ptodate/1/laa/5B_COMB.pep:*
3: /cgn2_6/ptodate/1/laa/6A_COMB.pep:*
5: /cgn2_6/ptodate/1/laa/6B_COMB.pep:*
5: /cgn2_6/ptodate/1/laa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodate/1/laa/PCTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-836-075A-10
US-08-635-886C-198
US-08-635-886C-232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-501-195-2
US-08-836-075A-46
US-08-836-075A-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US93-08638-9
PCT-US94-07088-23
PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-836-075A-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-904-686A-8
US-09-315-850-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-13552-2
                                                                                                                                                                                                                                                                                                                            156
1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-384-616-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-324-977-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             328717 seqs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                      - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                       US-09-491-146A-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

9100

                                                                                                                                                                                                                                                                                                                                                                                                                      Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                                                                                                               Perfect score:
                                                                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Result
```

Sequence 16, Appl Sequence 1, Applil Sequence 155, App Sequence 156, App Sequence 159, App Sequence 159, App Sequence 169, App Sequence 161, App Sequence 161, App Sequence 163, App Sequence 164, App Sequence 166, App Sequence 166, App Sequence 166, App Sequence 166, App Sequence 166, App Sequence 169, App Sequence 169, App Sequence 169, App Sequence 170, App Sequence 170, App	METHOD FOR MAKING THE SA	
1 US-07-681-701-16 3 US-08-078-2718-1 2 US-08-290-665A-155 2 US-08-290-665A-156 2 US-08-290-665A-156 2 US-08-290-665A-159 2 US-08-290-665A-159 2 US-08-290-665A-161 2 US-08-290-665A-161 2 US-08-290-665A-164 2 US-08-290-665A-165 12 US-08-290-665A-165 12 US-08-290-665A-165 12 US-08-290-665A-165 12 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167 2 US-08-290-665A-167	AL 1887	T NUMBER: 03063-0380 INFORMATION: -818-3700 18-3799 10
28 142 91.0 190 29 142 91.0 191 31 142 91.0 191 32 142 91.0 191 33 142 91.0 191 34 142 91.0 191 35 142 91.0 191 36 142 91.0 191 37 142 91.0 191 40 142 91.0 191 41 142 91.0 191 43 142 91.0 191 44 142 91.0 191 45 142 91.0 191	RESULT 1  Sequence 26, Application US/08921887; Sequence 26, Application US/08921887; Sequence 26, Application US/08921887; Sequence 26, Application US/08921887; Settle OF INVENTION: TITLE OF INVENTION: TITLE OF INVENTION: TITLE OF INVENTION: CORRESPONDENCE ADDRESS: CONTRY: USA COUNTRY: USA COUNTRY: USA COMPUTER: IBM PC COMPATIBLE COMPUTER: IBM PC COMPATIBLE COMPUTER: DATENTIN BATA: APPLICATION NUMBER: US/08/921, FILING DATE: CLASSIFICATION: NAME: WARREN, WILLIAM I. REGISTRATION NUMBER: 36,714	REFERENCE/DOCKET NUMBER: 0 TELECOMMUNICATION INFORMATION: 404-818-3700 TELEFAX: 404-818-3709 INFORMATION FOR SEQ ID NO: 26: 550UENCE CHARACTERISTICS: LENGTH: 28 amino acids TYPE: amino acids TYPANDEDNESS: not relevant TOPOLOGY: not relevant MOLECULE TYPE: protein HYPOTHERICAL: NO ANTI-SENSE: NO FRAGMENT TYPE: internal ORGANISM: HEPATILIS VIEUS US-08-921-887-26

0; Gaps

Length 28, Indels

100.0%; Score 156; DB 3; 100.0%; Pred. No. 1e-14; ive 0; Mismatches 0;

Query Match 100. Best Local Similarity 100. Matches 28; Conservative

```
SYNTHETIC POLYPEPTIDES BELONGING TO THE HEPATITIS C VIRUS (HCV) AND WHICH CAN BE USED ESPECIALLY FOR DETECTING THE LATTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 142; DB 4; Length 43;
Pred. No. 1.3e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
CLASSIFICATION NUMBER: US/09/020,846
FILING APPLICATION DATA:
APPLICATION NUMBER: 199 0.27015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION NUMBER: JP 9.027015
FILING DATE: 09-FEB-1996
ATTON NUMBER: US-FEB-1996
ATTON NUMBER: 25,258
REGISTRATION NUMBER: 25,258
REGISTRATION NUMBER: 25,258
TELEPHONE: (202) 672-5300
APPLICANT: CHIBA, Yukie
APPLICANT: YAGI, Shintaro
APPLICANT: HASEGAWA, AKIRA
APPLICANT: HASEGAWA, AKIRA
ATILE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                       COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2, Application US/08380160 Patent No. 6235284
                                                                                                                                                              STATE: D.C. N.W. STATE: D.C. S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET: P.O. BOX 19928
CITY: Alexandria
STATE: VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: DALBON, Pascal
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETI
TITLE OF INVENTION: HEPATITI
TITLE OF INVENTION: FOR DETE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDG
STREET: P.O. BOX 19928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 91.0%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-020-846-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                              (HCV) AND WHICH CAN BE USED ESPECIALLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                         APPLICANT: DALBON, PASCAL
APPLICANT: DALBON, PASCAL
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION: HEPATITIS C VIRUS (HCV) AND WHICH CAN BE
TITLE OF INVENTION: FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
ADDRESSE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
STATE: VA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.0%; Score 142; DB 3; Length 34; 92.9%; Pred. No. 9.9e-13; tive 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-00S/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 530
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AEBNT INFORMATION:
NAME: BETLIGG, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
TELEPHONE: (703)936-6400.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Human Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Sequence 36, Application US/09020846; Patent No. 6322965; GENERAL INFORMATION:
                                                                                                                                                                                                                              Sequence 6, Application US/08380160
Patent No. 6235284
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YAMAGUCHI, Kenjiro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNES: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FRAGMENT TYPE: N-terminal ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 92.9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-020-846-36
                                                                                                                                                                                                 US-08-380-160-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-380-160-6
ò
                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
```

Gaps ö

```
Sequence 1, Application US/08380160

Patent No. 6235284
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT:
JOLIVET, Michel
TITLE OF INVENTION:
SYNTHETIC POLYPEPTIDES BELONGING TO THE
TITLE OF INVENTION:
FOR DETECTING THE LATTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
Patent No. 6576240
                                                                                                                                                                                                                                               Query Match

91.0%; Score 142; DB 4; Length 44;
Best Local Similarity 92.9%; Pred. No. 1.3e-12;
Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION YUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                     4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28682
                                                                       TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Human Hepatitis C Virus STRAIN: H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOREZ/PORTION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: OLIFF & BERRIDGE
STREET: P.O. BOX 19928
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 45 amino acids
                                                                                                                                                                                      ; RELEVANT RESIDUES: 2 TO 45 US-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: JOLIVET, MICHEL
APPLICANT: PENIN, FRANCOIS
APPLICANT: DALBON, PASCAL
APPLICANT: LALBAN'ERE, LADAVIERE, LADAVIERE, LACOUX, XAVIER
APPLICANT: LACOUX, XAVIER
TITLE OF INVENTION: ANTIGENIC STRUCTURAL PEPTIDE, ANTIGENIC AND IMMUNOGENIC
TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND
FILLE REFERENCE: 103959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
      COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 142; DB 3; I
Pred. No. 1.3e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION NUMBER: US/09/389,756
CURRENT FILING DATE: 1999-09-07
EARLIER APPLICATION NUMBER: PCT/FR98/00442
EARLIER FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 11
SCPTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ANTI-SENSE: NO
FRACMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Human Hepatitis C Virus
STRAIN: H77
                                                                                                                                            CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,471
FILING DATE: 06-MAY-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequence 1, Application US/09389756
                                                                                                                                                                                                                                                         NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-6400
TELEFAX: (703)836-2787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Hepatitis C virus PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acids
TYPE: amino acids
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; OTHER INFORMATION: 7; OTHER INFORMATION: p; CTHER INFORMATION: the US-08-380-160-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: Peptide LOCATION: 1..44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SENERAL INFORMATION:
                                                                                                                              FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :ULT 5
-09-389-756-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
No. 5639594el Linear and Branched Peptides Effective
in Diagnosing and Detecting No. 5639594-A, No. 5639594-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 3, Application US/08530550
Patent No. 5736321
GENERAL INFORMATION:
APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFWARE: PATENTIN RELEASE #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,947
FILING DATE: 19930628
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 946,054
FILING DATE: 15-SEP-1993
ATTORREY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWHARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%; Score 142; DB 1;
92.9%; Pred. No. 1.8e-12;
tive 0; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PC-DOS/MS-DOS
              TITLE OF INVENTION: NO. 563955
TITLE OF INVENTION: 1n Diagnos
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34045
REFERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
SOFTWARE: Patentin Release #
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (516)273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                ZIP: 11788
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                        Hauppauge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: 25 LUC
                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Wang, Chang Yi
APPLICANT: Hoseln, Barbara H
TITLE OF INVENTION: No. 5582968el Branched Hybrid and Cluster
TITLE OF INVENTION: No. 5582968-B Hepatitis
TITLE OF INVENTION: No. 5582968-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                       /note= "N-terminal sequence of the protein of the nucleocapside or CORE protein of the human hepatitis C virus"
                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                     Ouery Match
91.0%; Score 142; DB 3; Length 45;
Best Local Similarity 92.9%; Pred. No. 1.3e-12;
Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142; DB 1; Length 61;
Pred. No. 1.8e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,054
FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
                                                                                                                                                                                                  1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                         5 PKPORKTKRNTNRRPODVKFPGGGGIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PKPQRKTKRNINRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 23, Application US/08083947
Patent No. 5639594
GENERAL INFORMATION
APPLICANT: Wang, Chang Yi
APPLICANT: Hosein, Barbara
                                                                                                                                                                                                                                                                                                                                            Sequence 9, Application US/07946054
Patent No. 5582968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 92.95
Best Local 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 516-273-1717
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
COCATION: 1..45
COTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
US-08-380-160-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CITY: Hauppauge
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                   US-07-946-054-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-083-947-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

ò

Gaps

```
Sequence 9, Application PC/TUS9308638
GENERAL INFORMATION:
GENERAL UNFORMATION:
TITLE OF INVENTION:
TITLE OF INVENTION:
FITLE OF INVENTION:
TITLE OF INVENTION:
Non-B Hepatitis
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSE:
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
CORRESPONDENCE ADDRESS:
                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 61;
                                                     Score 142; DB 1; Length 61;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142; DB 5;
Pred. No. 1.8e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Wang, Chang Yi
APPLICANT: Woseln, Barbara
TITLE OF INVENTION: Novel Linear And Branched
TITLE OF INVENTION: Peptides Effective In
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Diagnosing And Detecting
Non-A, Non-B Hepatitis
23
                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                        4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                     1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 23, Application PC/TUS9407088 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                      E: UNITED BIOMEDICAL INC. 25 Davids Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE: Floppy disk COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: M. LISA WILSON
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 905:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516_273-2828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.0%;
92.9%;
                                                         91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
                                                         Query Match
Best Local Similarity 92.99
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; TOPOLOGY: linear; MOLECULE TYPE: peptide pcr-us93-08638-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES:
    Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hauppauge
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PCT-US94-07088-23
                                                                                                                                                                                                                                                             PCT-US93-08638-9
;
US-08-262-037-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
                                                                                                                                                                                                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                            à
                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Chang Y1 Wang and Barbara Hosein
TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                      Length 61
                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                     Score 142; DB 1; Len'
---- NO. 1.8e-12;
2;
                                                                                                                                                                                                                                                                                                                             0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29,323
R: 1151-4043 US3
                                                                                                                                                                                                                                                                                                                                                                1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                           4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC COMPATIBLE
OPERATING SYGTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPREECT 5.1
SOFTWARE: APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 424
PRIOR APPLICATION ATA:
APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/667,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/510,153
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-APF11-1990
ATTORNEY/AGENT INFORMATION:
ANAMELY AREAS ANAMELY AT 11100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 26, Application US/08262037
Patent No. 5747239
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: MORGAN & FINNEGAN STREET: 345 PARK AVE. CITY: NEW YORK STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: MARIA C. H. Lin
REGISTRATION NUMBER: 29,323
REPERDICE/POCKET UNBER: 11:
TELECOMMUNICATION:
TRIEPHONE: 212-758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
                          REGISTRATION NUMBER: 34,045
REPERENCE/DOCKET NUMBER: 20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative 0
                                                                              TELEPHONE: (516)273-2828
TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHRACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 61 amino acids
                                                                                                                                                                    61 amino acids
                                                                                                                                                                                                                      MOLECULE TYPE: peptide US-08-530-550-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Amino acid
                                                                                                                                                                                     TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: Amino
STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: NI
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-262-037-26
                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                                                        à
```

Gaps

```
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: STUVYER, LIEVER
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOST:
TITLE OF INVENTION: AGENTS
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                               Query Match
91.0%; Score 142; DB 5; Length 61;
Best Local Similarity 92.9%; Pred. No. 1.8e-12;
Matches 26; Conservative 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: ISM FC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 142; DB 3; I
Pred. No. 2.2e-12;
0; Mismatches 2;
        REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4118PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEPRA: (212)751-6849
INFORMATION FOR SEQ ID NO: 3: SEQUENCE CHRARCTERISTICS:
LENGTH: 61 amino acids
                                                                                                                                                                                                                                                                                                                                                                                   1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                           4 PKPQKKTKKNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAPLICATION NUMBER: US/08/836,075A
PRIOR APPLICATION DATE: 21 Apr 1997
PRIOR APPLICATION DATE: APPLICATION NUMBER: PCT/EP95/04155
FILING DATE: 23 Oct 1995
PRIOR APPLICATION DATA: APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
PRIOR APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INNS: 004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Sequence 10, Application US/08836075A
; Patent No. 6180768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: KAMMERER, PATRICIA A. REGISTRATION NUMBER: 29,775 REFERENCE/DOCKET NUMBER: IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 91.0
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P.O. BOX 4433
                                                                                                                                                                                                                                  ; MOLECULE TYPE: peptide
PCT-US95-13660-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
US-08-836-075A-10
                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-836-075A-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: Peptides Effective for
TITLE OF INVENTION: Diagnosis and Detection of Hepatitis C Infection
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142; DB 5;
Pred. No. 1.8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                          COMPTRY: NEW LOLA
COMPUTER: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-TOS/MS-DOS
SOFTWARE: WORDERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07088
FILING DATE: 22-JUNE-1994
CLASSIFICATION: 108/083,947
FILING DATE: 28-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: LIN, MARIA C.H.
REGISTRATION NUMBER: 29323
REFERENCE/POCKET NUMBER: 1151-4101PC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4809
TELEPHONE: (212)758-4809
TELEPHONE: (212)758-4809
TELEPHONE: (212)758-4809
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1151-4101PC1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: Morgan & Finnegan, L.L.P
STREET: 345 Park Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 3, Application PC/TUS9513660 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/333,573
FILING DATE: 01 NOVEMBER 1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: (212)751-564-5
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
               ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 91.0%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; TOPOLOGY: 11near
; MOLECULE TYPE: peptide
PCT-US94-07088-23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PCT-US95-13660-3
```

ò g

9 P Search completed: August 7, 2003, 11:23:50 Job time : 10.6364 secs

us-09-491-146a-26.rapb

OM protein - protein search, using sw model

(without alignments) 231.506 Million cell updates/sec August 7, 2003, 11:20:11; Search time 14.3636 Seconds Run on:

US-09-491-146A-26

Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

451899 segs, 118759770 residues Searched: tal number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Published\_Applications\_AA:\* Database :

/cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\* /ptcdata/1/pubpaa/US09B\_PUBCOMB.pep: /ptcdata/1/pubpaa/US09C\_PUBCOMB.pep: /ptcdata/1/pubpaa/US09\_NEW\_PUB.pep: /ggn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG6\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG6\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG0\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG0\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG0\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG0\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG0\_NEW\_PUB.pep:\*
/cgn2\_6/ptodata/1/pubpaa/USG0\_NEW\_PUB.pep:\* /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep 16: 17: 18:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

	Description	Sequence 1, Appli	Sequence 10, Appl	Sequence 1, Appli	8, A	Sequence 77, Appl	Sequence 14, Appl	Sequence 78, Appl	Sequence 46, Appl	Sequence 60, Appl	152,		42,	Sequence 44, Appl	Sequence 42, Appl	Sequence 44, Appl
	ΠD	US-10-367-677-1	US-09-851-138-10	US-09-758-308-1	US-09-756-875-8	US-09-921-397-77	US-09-851-138-14	US-09-921-397-78	US-09-851-138-46	US-09-851-138-60	US-09-899-046-152	US-09-878-281-152	US-09-899-046-42	US-09-899-046-44	US-09-878-281-42	US-09-878-281-44
	,	15	10	6	10	10	10	10	10	10	1	11	11	11	11	11
	Query Match Length DB	44	74	9.1	46	103	108	113	137	138	166	166	169	169	169	169
æ	Query	91.0	91.0	91.0	91.0	91.0	0.16	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91	91.0
	Score	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142
	Result No.		2	. ~	4		) V		- cc	σ	,0	-	12		14	15

Length 44;

DB 15;

Score 142;

91.0%;

Query Match

7	Sednence 7, Appli				Sequence 18, Appl	ò	52,	24, 7	Sequence 144, App		52,		Sequence 144, App	Sequence 50, Appl	20,	Sequence 50, Appl	50,	23,		Sequence 40, Appl	4	o,	Sequence 1, Appli	Sequence 20, Appl	Sequence 3, Appli	Sequence 1, Appli	20	cvi	Sequence 2, Appli
-929	US-10-104-966	US-09-194-9	US-09-851-	9-851-1	US-09-851-	-668-60	11 US-09-899-046-52	S	11 US-09-899-046-144	6	978-	us-	۹		US-09-8			10 US-09-941-611-23			9 US-09-742-659-4	US-09-952			11 US-09-891-894-3		15 US-10-259-275-20	ns	11 US-09-995-937-2
182 1	182	191	318	319	319	319	319	319	319	319	319	319	319	809	808	608	809	2894	2894	2985	3011	3011	3011	3011	3011	3011	3011	3012	3012
91.0	91.0	91.0	0.16	0.16		0.16	91.0	0.16	0.16	91.0	91.0	91.0	91.0	0.16	0.16	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0	91.0		91.0
142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142	142		142				142	142	142	142
16	17	<u>.</u>	0	500	2 6	100	1.0	) V	20	26	27	80	0	í.	3 6	100	. "	, e		3 0	3.7	30	90	0 0	. 4	4.2	4.4	44	4.5

## ALIGNMENTS

```
Sequence 1, Application US/10367677

Squence 1, Application US/10367677

Squence 1, Application WS. US2030118604A1

GENERAL INFORMATION:

APPLICANT: JOLIVET, MICHEL

APPLICANT: JOLIVET, MICHEL

APPLICANT: LACOUX, XAVIER

APPLICANT: LACOUX, XAVIER

TITLE OF INVENTION: COMPOUNDS, AND USES FOR DETECTING, PREVENTING AND TITLE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND FILE OF INVENTION: TREATING AND WIFECTION

FILE OF INVENTION: TREATING AND USES FOR DETECTING, PREVENTING AND FILE REFERENCE: 103959

CURRENT APPLICATION NUMBER: US/10/367,677

CURRENT FILING DATE: 2003-02-19

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/FR98/00442

PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-03-05

NUMBER OF SED ID NOS: 11

SGOTWARE: PATENTING DATE: 2.1

SEQ ID NO 1

FENCENTIANG DATE: SERVICE APPLICATION NUMBER: PTON NUMBER OF SED ID NOS: 11

FENCENTIANG DATE: SERVICE APPLICATION NUMBER: PATENT NUMBER OF SED ID NOS: 11

FENCENTIANG DATE: SERVICE APPLICATION NUMBER: PATENT NUMBER OF SED ID NOS: 11

FENCENTIANG DATE: SERVICE APPLICATION NUMBER: PATENT NUMBER OF SED ID NOS: 11

FENCENTIANG DATE: SERVICE APPLICATION NUMBER: PATENT NUMBER: PATENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS: Ogata, N. et al.
TITLE: Nucleotide Sequence and Mutation Rate of the H Strain
TITLE: of Hepatitis Virus
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RELEVANT RESIDUES: 2 TO 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAGES: 3392-3396
DATE: 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-10-367-677-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 44
US-10-367-677-1
```

```
Gaps
                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 97;
                                                                                                                                                                                                                     Length 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/756,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 142; DB 10;
Pred. No. 5.7e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: PIKE, IAN
TITLE OF INVENTION: HEPATITIS C VIRUS PEPTIDES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                 Score 142; DE
Pred. No. 5.3e
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ...UNKESSE: ...UNKESS:
STREET: 555 13th Street, N. W.
CITY: Washington
                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 PKPQRKPNRNTNRRPQDVKFPGGGGIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PCT/GB93/00410
                           60/092,339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US 08/259,721
                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 8, Application US/09756875 Patent No. US20020150990A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: Floppy disk COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/Me-SOFTWARE: Patent
       2001-01-10
                                                                                Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: ERNST, BARBARA G. 377
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 181
TELECOMMUNICATION INFORMATION:
                                         1999-07-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/FILING DATE: 29-AUG-1994
APPLICATION NUMBER: PCT/GB
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91.0%;
92.9%;
                                                                                                                                                                                                                 Query Match
Best Local Similarity 92.9%;
Matches 26; Conservative
   CURRENT FILING DATE: 2001-01
PRIOR APPLICATION NUMBER: 6
PRIOR FILING DATE: 1999-07-
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version
SEQ ID NO 1
LENGTH: 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (202)783-6040
                                                                                                                                      ; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-758-308-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 91.0°
Best Local Similarity 92.9°
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: unknown
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D. C.
: U. S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIRUS
TITLE OF INVENTION: PROTEINS
                                                                                                                                                                                                                                                          STUTYPER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                         Gaps
                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
 Pred. No. 2.5e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match
91.0%; Score 142; DB 10;
Best Local Similarity 92.9%; Pred. No. 4.3e-12;
Matches 26; Conservative 0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: ARNOLD, WHITE & DURKEE STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 10:
                                                         1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                               1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: peptide
;
US-09-851-138-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILE REFERENCE: 14114.0349U2
CURRENT APPLICATION NUMBER: US/09/758,308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: KAMMERER, PATRICIA A.
                                                                                                                                                                   5-09-851-138-10
Sequence 10, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Application US/09758308
Patent No. US20020090607A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 74 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                        AGENTS
92.98;
                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
US-09-758-308-1
                                                           ò
                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
```

```
STUTYVER, LIEVEN
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                    polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 113;
                                                         Length 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 142; DB 10;
Pred. No. 6.7e-12;
0; Mismatches 2;
                                                         Score 142; DB 10;
Pred. No. 6.4e-12;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: 08/836,075
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 94870166.9
                                                                                                                                                                                                                                                                                         Sequence 78, Application US/09921397

Patent No. US20020151484A1

GENERAL INFORMATION:
TITLE OF INVENTION: SID nucleic acids and polyTITLE OF INVENTION: pathogenic strain of the TITLE OF INVENTION: applications thereof FILE REFERENCE: B4809A JAZ

CURRENT APPLICATION NUMBER: US/09/921,397

CURRENT FILING DATE: 2001-08-02

PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 PKPORKPNRNTNRRPODVKFPGGGOIVG 28
                                                                                                                                                 1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
; SEQUENCE DESCRIPTION: SEQ ID NO: 14: US-09-851-138-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 46, Application US/09851138
Publication No. US20020183508A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91.0%;
ilarity 92.9%;
Conservative
                                                                  91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: TEXAS
COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQ ID NOS: 156
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 78
                                                                  Query Match
Best Local Similarity 92.93
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-921-397-78
                                                                                                                                                                                                                                                                                 US-09-921-397-78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                              polypeptides selected from
the hepatitis C virus and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION NUMBER: 08/836,075
FILING DATE: CUNKNOWN>
APPLICATION NUMBER: EP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142; DB 10;
Pred. No. 6.1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
           US-09-921-397-77

Sequence 77, Application US/09921397

Sequence 77, Application US/09921397

Patent No. US20020151484A1

GENERAL INFORMATION:

APPLICANT: HYBRIGENICS

TITLE OF INVENTION: SID nucleic acids and pol

TITLE OF INVENTION: pathogenic strain of the

TITLE OF INVENTION: applications thereof

FILE REFERENCE: B4809A - JAZ

CURRENT APPLICATION UNDMER: US/09/921.397

CURRENT FILING DATE: 2001-08-02

PRIOR FILING DATE: 2000-08-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 14, Application US/09851138 Publication No. US20020183508Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 108 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 91.0%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Hepatitis C virus 9-09-921-397-77
                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
ZIP: 77210-4433
                                                                                                                                                                                                                                                                                            NUMBER OF SEQ ID NOS: 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09-851-138-14
                                                                                                                                                                                                                                                                                                                                                           LENGTH: 103
                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy
NUMBER OF SEQUENCES: 270
            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 166;
              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
          .;
?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 142; DB 11;
Pred. No. 1e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 11;
          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Mismatches
                                                1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                             5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.0%; Score 142;
                                                                                                                                                                                    ; Sequence 152, Application US/09899046; Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                          PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 152, Application US/09878281 Publication No. US20030032005A1 GENERAL INFORMATION:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
          ö
                                                                                                                                                                                                                                                                                                                                                                     IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.0%;
                                                                                                                                                                                                                                                                                                                                                  Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 166 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 92.99
Matches .26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             166 amino acids
        Conservative
                                                                                                                                                                                                                                                                TITLE OF INVENTION: New TITLE OF INVENTION: GENC NUMBER OF SEQUENCES: 270 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: protein US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       inear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE:
US-09-878-281-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                 US-09-899-046-152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-878-281-152
        26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
        Matches
                                                ó
                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOft Word 6.0 / ASCII text output
CURRENT APPLICATION NUMBER: US/09/851,138
FILING DATE: 09-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: BP 94870166.9
FILING DATE: CURNOWN>
APPLICATION NUMBER: PP 94870166.9
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 22 9.3 Jun 1995
                                                                                                                                                                                                                                                                                                                           Length 137;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 138;
                                                                                                                                                                                                                                                                                                                      Score 142; DB 10; Length 1
Pred. No. 8.2e-12;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 142; DB 10;
Pred. No. 8.3e-12;
FILING DATE: 21 Oct 1994
APPLICATION NUMBER: EP 95870076.7
FILING DATE: 28 Jun 1995
ATTORNEY/AGENT INFORMATION:
                                                                             NAMÉ: KAMMEREK, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: ARNOLD, WHITE & DURKEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: KAMMERER, PATRICIA A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: INNS:004
INFORMATION FOR SEQ ID NO: 60:
                                                                                                                                                                                                                                                                                                                                                                                                    1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-851-138-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 60, Application US/09851138 Publication No. US20020183508A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 138 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: MAERTENS, GEERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STUYVER, LIEVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: P.O. BOX 4433
                                                                                                                                                                                                                                                                                                                    91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                      TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                              US-09-851-138-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-851-138-60
                                                                                                                                                                                                                                                                                                                        Query Match
```

õ g

```
Publication No. Control of the contr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 169;
Score 142; DB 11; Length 169;
Pred. No. 1e-11;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 42, Application US/09878281

Sequence 42, Application US/09878281

Sublication No. US20030032065A1

Fublication No. US20030032065A1

SENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Genctypes for diagnosis, prophylaxis a TITLE OF INVENTION: genctypes for diagnosis, prophylaxis a UNMER OF SEQUENCES: 270

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
COMPUTER: PLACE FORM:
COMPUTER:
COMPUTER: PLACE FORM:
COMPUTER: PLACE FORM:
COMPUTER:
COMPUTER:
COMPUTER: PLACE FORM:
COMPUTER:

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 142; DB
Pred. No. 1e-11
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPORKPNRNTNRRPODVKFPGGGQIVG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                    5 PKPQRKTKRNTNRRPQDVKFPGGGOIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-878-281-44; Sequence 44, Application US/09878281; Publication No. US20030032005A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91.0%;
92.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                       Ouery Match 91.0%;
Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   , MOLECULE TYPE: protein US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-878-281-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-878-281-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ô
                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
TITLE OF INVENTION: New sequences of hepatitis C virus
TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
TITLE OF SEQUENCES: 270
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy, 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 11; Length 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 42, Application US/09899046;
Sequence 42, Application US/09899046;
Publication No. US20030008274A1
GENERAL INFORMATION:
APPLICAMY:
APPLICAMY:
TITLE OF INVENTION: Genotypes for diagnosis, prophylaxis;
TITLE OF INVENTION: Genotypes for diagnosis, prophylaxis;
COMPUTER READALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Competible
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
COMPUTER TANNER TO THE TO THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: FIOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARRE: Patentin Release #1.0, Version #1.25 (EPO)
CURENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,046
FILING DATE:
                                                                                                       Indels
                                                                                                       ;
;
                                                             Pred. No. 1e-11;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.0%; Score 142; DE
92.9%; Pred. No. 1e-1
Live 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PKPQRKTKRNTNRRPQDVKFPGGGQIVG 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                 1 PKPQRKPNRNTNRRPQDVKFPGGGQIVG 28
                                                                                                                                                                                                                                                APPLICATION NUMBER: US/09/899,046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 44, Application US/09999046 vublication No. US20030008274A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
                                                             Best Local Similarity 92.9%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 169 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 91.0
Best Local Similarity 92.9
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , MOLECULE TYPE: protein US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
US-09-899-046-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-899-046-42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      αd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
```

ţ

Ouery Match 91.0%; Score 142; DB 11; Length 169; Best Local Similarity 92.9%; Pred. No. 1e-11; Matches 26; Conservative 0; Mismatches 2; Indels (

0

Gaps

ö

g ò

Search completed: August 7, 2003, 12:01:13 Job time : 15.3636 secs